Parameter estimation for an affine two factor model

MÁTYÁS BARCZY*, LEIF DÖRING, ZENGHU LI, GYULA PAP

Abstract

For an affine two factor model, we study the asymptotic properties of the maximum likelihood and least squares estimators of some appearing parameters in the so-called subcritical (ergodic) case based on continuous time observations. We prove strong consistency and asymptotic normality of the estimators in question.

1 Introduction

We consider the following 2-dimensional affine process (affine two factor model)

$$\begin{cases} \mathrm{d}Y_t = (a - bY_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}L_t, & t \geqslant 0, \\ \mathrm{d}X_t = (m - \theta X_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}B_t, & t \geqslant 0, \end{cases} \tag{1.1}$$

where $a > 0$, $b, \theta, m \in \mathbb{R}$, $(L_t)_{t\geqslant 0}$ and $(B_t)_{t\geqslant 0}$ are independent standard Wiener processes. Note that the process $(Y_t)_{t\geqslant 0}$ given by the first SDE of (1.1) is the so-called Cox-Ingersol-Ross (CIR) process which is a continuous state branching process with branching mechanism $bz + z^2/2$, $z \geqslant 0$, and with immigration mechanism az, $z \geqslant 0$. Chen and Joslin [8] have found several applications of the model (1.1) in financial mathematics, see their equations (25) and (26).

The process (Y, X) given by (1.1) is a special affine diffusion process. The set of affine processes contains a large class of important Markov processes such as continuous state branching processes and Orstein-Uhlenbeck processes. Further, a lot of models in financial mathematics are affine such as the Heston model [11], the model of Barndorff-Nielsen and Shephard [3] or the model due to Carr and Wu [7]. A precise mathematical formulation and a complete characterization of regular affine processes are due to Duffie et al. [10].

This article is devoted to estimate the parameters m and θ from some continuously observed real data set. To the best knowledge of the authors the parameter estimation problem for multi-dimensional affine processes has not been tackled so far. Since affine processes are frequently used in financial mathematics, the question of parameter estimation for them is of high importance. In Barczy et al. [1] we started the discussion with the simple non-trivial two-dimensional affine diffusion process given by (1.1) in the so called critical case (for the definition of criticality, see Section 2). We examined (conditional) least squares estimator of (m, θ) from some discretely observed real data set. In this paper we deal with the same model but in the so-called subcritical (ergodic) case (using a continuously observed real data set). In case of the one-dimensional CIR process Y, the parameter

$*$ Corresponding author

2010 Mathematics Subject Classifications: 62F12, 60J25.

Key words and phrases: affine process, maximum likelihood estimator, least squares estimator.

M. Barczy, Z. Li and G. Pap have been supported by the Hungarian Chinese Intergovernmental S & T Cooperation Programme for 2011-2013 under Grant No. 10-1-2011-0079. M. Barczy and G. Pap have been partially supported by the Hungarian Scientific Research Fund under Grant No. OTKA T-079128. Z. Li has been partially supported by NSFC under Grant No. 11131003 and 973 Program under Grant No. 2011CB808001.

estimation of a and b goes back to Overbeck and Rydén [21], Overbeck [22], and see also the very recent papers of Ben Alaya and Kebaier [4], [5]. Finally, we note that Li and Ma [18] started to investigate the asymptotic behaviour of the (weighted) conditional least squares estimators of the drift parameters for a CIR model driven by a stable noise (they call it a stable CIR model) from some discretely observed real data set.

For studying the asymptotic behaviour of the maximum likelihood and least squares estimators of (m, θ) in the subcritical (ergodic) case, one first needs to examine the question of existence of a unique stationary distribution and ergodicity for the model given by (1.1). In a companion paper Barczy et al. [2] we solved this problem in a more general setup by replacing the CIR process $(Y_t)_{t\geqslant 0}$ in the first SDE of (1.1) by a so-called α-root process (stable CIR process) with $\alpha \in (1,2)$.

We give a brief overview of the structure of the paper. Section 2 is devoted to a preliminary discussion of the existence and uniqueness of a strong solution of the SDE (1.1), we also recall our results in Barczy et al. [2] on the existence of a unique stationary distribution and ergodicity for the affine process given by SDE (1.1), see Theorem 2.3. Further, we define and study criticality of the model (1.1), and we recall some limit theorems for continuous local martingales that will be used later on for studying the asymptotic behaviour of maximum likelihood and least squares estimators of (m, θ). In Sections 3 – 8 we study the asymptotic behavior of the maximum likelihood and least squares estimator of (m, θ) proving that the estimators are strongly consistent and asymptotically normal under appropriate conditions on the parameters. We note that in the critical case we obtained a different limit behaviour for the least squares estimator of (m, θ), see Barczy et al. [1, Theorem 3.2].

2 Preliminaires

Let $\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the sets of positive integers, non-negative integers, real numbers and non-negative real numbers, respectively. By $\|x\|$ and $\|A\|$ we denote the Euclidean norm of a vector $x \in \mathbb{R}^m$ and the induced matrix norm $\|A\| = \sup\{\|Ax\| : x \in \mathbb{R}^m,\ \|x\| = 1\}$ of a matrix $A \in \mathbb{R}^{n\times m}$, respectively.

Let $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P}\big)$ be a filtered probability space satisfying the usual conditions, i.e., $(\Omega, \mathcal{F}, \mathbb{P})$ is complete, the filtration $(\mathcal{F}_t)_{t\geqslant 0}$ is right-continuous and $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets in $\mathcal{F}$. Let $(L_t)_{t\geqslant 0}$ and $(B_t)_{t\geqslant 0}$ be independent standard $(\mathcal{F}_t)_{t\geqslant 0}$-Wiener processes.

The next proposition is about the existence and uniqueness of a strong solution of the SDE (1.1), see Proposition 2.1 with $\alpha = 2$ in Barczy et al. [2].

2.1 Proposition. *Let* (Y_0, X_0) *be any* $\mathcal{F}_0$*-measurable random vector such that* $\mathbb{P}(Y_0 \geqslant 0) = 1$. *Then, for all* $a \in \mathbb{R}_+$, $b, m, \theta \in \mathbb{R}$, *there is a (pathwise) unique strong solution* $(Y_t, X_t)_{t\geqslant 0}$ *of the SDE* (1.1) *such that* $\mathbb{P}(Y_t \geqslant 0,\ \forall\, t \geqslant 0) = 1$. *Further, we have*

$$Y_t = \mathrm{e}^{-b(t-s)}\left(Y_s + a\int_s^t \mathrm{e}^{-b(s-u)}\,\mathrm{d}u + \int_s^t \mathrm{e}^{-b(s-u)}\sqrt{Y_u}\,\mathrm{d}L_u\right), \qquad 0 \leqslant s \leqslant t, \tag{2.1}$$

and

$$X_t = \mathrm{e}^{-\theta(t-s)}\left(X_s + m\int_s^t \mathrm{e}^{-\theta(s-u)}\,\mathrm{d}u + \int_s^t \mathrm{e}^{-\theta(s-u)}\sqrt{Y_u}\,\mathrm{d}B_u\right), \qquad 0 \leqslant s \leqslant t. \tag{2.2}$$

2.2 Remark. (i) In Proposition 2.1 the assumption that (Y_0, X_0) is $\mathcal{F}_0$-measurable yields that (Y_0, X_0) is independent (of the increments) of $(L_t, B_t)_{t \geqslant 0}$.

(ii) In Proposition 2.1 it is the assumption $a \in \mathbb{R}_+$ which ensures $\mathbb{P}(Y_t \geqslant 0, \ \forall\, t \geqslant 0) = 1$.

(iii) The existence of a (pathwise) unique strong solution $(Y_t, X_t)_{t \geqslant 0}$ of the SDE (1.1) such that $\mathbb{P}(Y_t \geqslant 0, \ \forall\, t \geqslant 0) = 1$ follows also by a general result of Dawson and Li [9, Theorem 6.2]. □

In the sequel $\xrightarrow{\mathbb{P}}$ and $\xrightarrow{\mathcal{L}}$ will denote convergence in probability and in distribution, respectively.

The following result states the existence of a unique stationary distribution and the ergodicity for the affine process given by the SDE (1.1), see Theorems 3.1 and 3.2 with $\alpha = 2$ in Barczy et al. [2].

2.3 Theorem. *Let us consider the 2-dimensional affine model* (1.1) *with* $a > 0$, $b > 0$, $m \in \mathbb{R}$, $\theta > 0$, *and with a (random)* $\mathcal{F}_0$*-measurable initial value* (Y_0, X_0) *such that* $\mathbb{P}(Y_0 \geqslant 0) = 1$. *Then*

(i) $(Y_t, X_t) \xrightarrow{\mathcal{L}} (Y_\infty, X_\infty)$ *as* $t \to \infty$, *and the distribution of* (Y_∞, X_∞) *is given by*

$$\mathbb{E}\left(\mathrm{e}^{-\lambda_1 Y_\infty + i\lambda_2 X_\infty}\right) = \exp\left\{ -a \int_0^\infty v_s(\lambda_1, \lambda_2)\,\mathrm{d}s + i\frac{m}{\theta}\lambda_2 \right\}, \qquad (\lambda_1, \lambda_2) \in \mathbb{R}_+ \times \mathbb{R}, \tag{2.3}$$

where $v_t(\lambda_1, \lambda_2)$, $t \geqslant 0$, *is the unique non-negative solution of the (deterministic) differential equation*

$$\begin{cases} \frac{\partial v_t}{\partial t}(\lambda_1, \lambda_2) = -b v_t(\lambda_1, \lambda_2) - \frac{1}{2}(v_t(\lambda_1, \lambda_2))^2 + \frac{1}{2}\mathrm{e}^{-2\theta t}\lambda_2^2, & t \geqslant 0, \\ v_0(\lambda_1, \lambda_2) = \lambda_1. \end{cases} \tag{2.4}$$

(ii) *supposing that the random initial value* (Y_0, X_0) *has the same distribution as* (Y_∞, X_∞) *given in part (i), we have* $(Y_t, X_t)_{t \geqslant 0}$ *is strictly stationary.*

(iii) *for all Borel measurable functions* $f : \mathbb{R}^2 \to \mathbb{R}$ *such that* $\mathbb{E}\left(|f(Y_\infty, X_\infty)|\right) < \infty$, *we have*

$$\mathbb{P}\left(\lim_{T \to \infty} \frac{1}{T} \int_0^T f(Y_s, X_s)\,\mathrm{d}s = \mathbb{E}(f(Y_\infty, X_\infty)) \right) = 1, \tag{2.5}$$

where the distribution of (Y_∞, X_∞) *is given by* (2.3) *and* (2.4).

Moreover, the random variable (Y_∞, X_∞) *is absolutely continuous, the Laplace transform of* Y_∞ *takes the form*

$$\mathbb{E}(\mathrm{e}^{-\lambda_1 Y_\infty}) = \left(1 + \frac{\lambda_1}{2b}\right)^{-2a}, \qquad \lambda_1 \in \mathbb{R}_+, \tag{2.6}$$

i.e., Y_∞ *has Gamma distribution with parameters* $2a$ *and* $2b$, *all the (mixed) moments of* (Y_∞, X_∞) *of any order are finite, i.e.,* $\mathbb{E}(Y_\infty^n |X_\infty|^p) < \infty$ *for all* $n, p \in \mathbb{Z}_+$, *and especially,*

$$\mathbb{E}(Y_\infty) = \frac{a}{b}, \qquad \mathbb{E}(X_\infty) = \frac{m}{\theta},$$

$$\mathbb{E}(Y_\infty^2) = \frac{a(2a+1)}{2b^2}, \qquad \mathbb{E}(Y_\infty X_\infty) = \frac{ma}{\theta b}, \qquad \mathbb{E}(X_\infty^2) = \frac{a\theta + 2bm^2}{2b\theta^2},$$

$$\mathbb{E}(Y_\infty X_\infty^2) = \frac{a}{(b+2\theta)2b^2\theta^2}\big(\theta(ab + 2a\theta + \theta) + 2m^2 b(2\theta + b)\big).$$

In what follows we define and study criticality of the affine process given by the SDE (1.1). The next definition of criticality is based on Proposition 3.2 in Barczy et al. [1].

2.4 Definition. *Let* $(Y_t, X_t)_{t\in\mathbb{R}_+}$ *be an affine diffusion process given by the SDE* (1.1) *with an* $\mathcal{F}_0$*-measurable initial value* (Y_0, X_0) *satisfying* $\mathbb{P}(Y_0 \geqslant 0) = 1$. *We call* $(Y_t, X_t)_{t\in\mathbb{R}_+}$ *subcritical, critical or supercritical if the spectral radius of the matrix*

$$\begin{pmatrix} \mathrm{e}^{-bt} & 0 \\ 0 & \mathrm{e}^{-\theta t} \end{pmatrix}$$

is less than 1, *equal to* 1 *or greater than* 1, *respectively.*

Note that, since the spectral radius of the matrix given in Definition 2.4 is $\max(\mathrm{e}^{-bt}, \mathrm{e}^{-\theta t})$, the affine process given in Definition 2.4 is

$$\begin{array}{rl} \text{subcritical} & \text{if } \ b > 0 \ \text{ and } \ \theta > 0, \\[1ex] \text{critical} & \text{if } \ b = 0, \ \theta \geqslant 0 \ \text{ or } \ b \geqslant 0, \ \theta = 0, \\[1ex] \text{supercritical} & \text{if } \ b < 0 \ \text{ or } \ \theta < 0. \end{array}$$

Definition 2.4 of criticality is in accordance with the corresponding definition for one-dimensional continuous state branching processes, see, e.g., Li [17, page 58].

In all what follows we will suppose that we have continuous time observations for the process (Y, X), i.e., $(Y_t, X_t)_{t\in[0,T]}$ can be observed for some $T > 0$, and our aim is to deal with parameter estimation of θ and (θ, m), respectively, provided that the parameters $a > 0$ and $b \in \mathbb{R}$ are supposed to be known.

In general, parameter estimation for subcritical (also called ergodic) models has a long history, see, e.g., the monographs of Kutoyants [16] and Bishwall [6]. To give some examples, we mention two models that are somewhat related to (1.1) and parameter estimation of the appearing parameters based on continuous time observations has been considered. In the introduction we have already mentioned the one-dimensional CIR process given by the first SDE of (1.1),

$$\mathrm{d}Y_t = (a - bY_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}L_t, \qquad t \geqslant 0,$$

for which the parameter estimation of a and b goes back to Overbeck and Rydén [21], Overbeck [22], and see also the very recent papers of Ben Alaya and Kebaier [4], [5]. For parameter estimation for α-stable CIR processes with $\alpha \in (1, 2]$, see Li and Ma [18]. The second model that we mention is the so-called Ornstein-Uhlenbeck process driven by α-stable Lévy motions, i.e.,

$$\mathrm{d}U_t = (m - \theta U_t)\,\mathrm{d}t + \mathrm{d}Z_t, \qquad t \geqslant 0,$$

where $\theta > 0, m \neq 0$, and $(Z_t)_{t\geqslant 0}$ is an α-stable Lévy motion with $\alpha \in (1, 2)$. For this model Hu and Long investigated the question of parameter estimation, see [12], [13] and [14].

In what follows we recall some limit theorems for continuous local martingales. We will use these limit theorems in the sequal for studying the asymptotic behaviour of different kinds of estimators for (m, θ). First we recall a strong law of large numbers for continuous local martingales, see, e.g., Liptser and Shiryaev [20, Lemma 17.4].

2.5 Theorem. *Let* $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P}\big)$ *be a filtered probability space satisfying the usual conditions. Let* $(M_t)_{t\geqslant 0}$ *be a continuous local martingale with respect to the filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ *started from* 0. *Let* $(\xi_t)_{t\geqslant 0}$ *be a progressively measurable process such that*

$$\mathbb{P}\left(\int_0^t (\xi_u)^2 \, \mathrm{d}\langle M\rangle_u < \infty\right) = 1, \qquad t \geqslant 0,$$

and

$$\mathbb{P}\left(\lim_{t\to\infty} \int_0^t (\xi_u)^2 \, \mathrm{d}\langle M\rangle_u = \infty\right) = 1, \tag{2.7}$$

where $(\langle M\rangle_t)_{t\geqslant 0}$ *denotes the quadratic variation process of* M. *Then*

$$\mathbb{P}\left(\lim_{t\to\infty} \frac{\int_0^t \xi_u \, \mathrm{d}M_u}{\int_0^t (\xi_u)^2 \, \mathrm{d}\langle M\rangle_u} = 0\right) = 1. \tag{2.8}$$

In case of $M_t = B_t$, $t \geqslant 0$, *where* $(B_t)_{t\geqslant 0}$ *is a standard Wiener process, the progressive measurability of* $(\xi_t)_{t\geqslant 0}$ *can be relaxed to measurability and adaptedness to the filtration* $(\mathcal{F}_t)_{t\geqslant 0}$.

The next theorem is about the asymptotic behaviour of continuous multivariate local martingales.

2.6 Theorem. (van Zanten [25, Theorem 4.1]) *Let* $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P}\big)$ *be a filtered probability space satisfying the usual conditions. Let* $(M_t)_{t\geqslant 0}$ *be a* d*-dimensional continuous local martingale with respect to the filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ *started from* 0. *Suppose that there exists a function* $Q : [0,\infty) \to \mathbb{R}^{d\times d}$ *such that* $Q(t)$ *is a non-random, invertible matrix for all* $t \geqslant 0$, $\lim_{t\to\infty} \|Q(t)\| = 0$ *and*

$$Q(t)\langle M\rangle_t Q(t)^\top \xrightarrow{\mathbb{P}} \eta\eta^\top \quad \text{as} \quad t \to \infty,$$

where η *is a* $d\times d$ *random matrix defined on* $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$. *Then, for each* $\mathbb{R}^k$*-valued random variable* V *defined on* $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$, *it holds that*

$$(Q(t)M_t, V) \xrightarrow{\mathcal{L}} (\eta Z, V) \qquad \text{as} \qquad t \to \infty,$$

where Z *is a* d*-dimensional standard normally distributed random variable independent of* (η, V).

3 Existence and uniqueness of maximum likelihood estimator

Let $\mathbb{P}_{(\theta,m)}$ denote the probability measure on $(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+ \times \mathbb{R}), \mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+ \times \mathbb{R})))$ induced by the process $(Y_t, X_t)_{t\geqslant 0}$ corresponding to the parameters (a, b, θ, m) and initial value (Y_0, X_0). Note that we do not denote the dependence on a and b, and on the initial value (Y_0, X_0), since in what follows in a given statement all the processes that we consider will have the same parameters a and b and the same initial value. Here we suppose that the space $(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+ \times \mathbb{R}), \mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+ \times \mathbb{R})))$ is endowed with the natural filtration $(\mathcal{A}_t)_{t\geqslant 0}$, given by $\mathcal{A}_t := \varphi_t^{-1}(\mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+ \times \mathbb{R})))$, where $\varphi_t : \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R}) \to \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R})$ is the mapping $\varphi_t(f)(s) := f(t\wedge s)$, $s \geqslant 0$, and $\mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R}))$ denotes the Borel σ-algebra on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+ \times \mathbb{R})$. For all $T > 0$, let $\mathbb{P}_{(\theta,m),T} := \mathbb{P}_{(\theta,m)}|_{\mathcal{A}_T}$ be the restriction of $\mathbb{P}_{(\theta,m)}$ to $\mathcal{A}_T$.

3.1 Lemma. *Let* $a \geqslant 1/2$, $b, \theta, m \in \mathbb{R}$, $T > 0$, *and suppose that* $\mathbb{P}(Y_0 > 0) = 1$. *Let* $\mathbb{P}_{(\theta,m)}$ *and* $\mathbb{P}_{(0,0)}$ *denote the probability measures induced by the unique strong solutions of the SDE* (1.1) *corresponding to the parameters* (a, b, θ, m) *and* $(a, b, 0, 0)$ *with the same initial value* (Y_0, X_0), *respectively. Then* $\mathbb{P}_{(\theta,m),T}$ *and* $\mathbb{P}_{(0,0),T}$ *are absolutely continuous with respect to each other, and the Radon-Nykodim derivative of* $\mathbb{P}_{(\theta,m),T}$ *with respect to* $\mathbb{P}_{(0,0),T}$ *(so called likelihood ratio) takes the form*

$$L_T^{(\theta,m),(0,0)}((Y_s, X_s)_{s\in[0,T]}) = \exp\left\{ \int_0^T \frac{(m - \theta X_s)}{Y_s} \,\mathrm{d}X_s - \frac{1}{2}\int_0^T \frac{(m-\theta X_s)^2}{Y_s}\,\mathrm{d}s \right\},$$

where $(Y_t, X_t)_{t\geqslant 0}$ *denotes the unique strong solution of the SDE* (1.1) *corresponding to the parameters* (a, b, θ, m) *and the initial value* (Y_0, X_0). *The process*

$$\left(L_T^{(\theta,m),(0,0)}((Y_s, X_s)_{s\in[0,T]})\right)_{T\geqslant 0} \tag{3.1}$$

is an $(\mathcal{F}_T)_{T\geqslant 0}$*-martingale (where the filtration* $(\mathcal{F}_T)_{T\geqslant 0}$ *is introduced before Proposition 2.1).*

Proof. First note that the SDE (1.1) can be written in the form:

$$\mathrm{d}\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \left[\begin{pmatrix} -b & 0 \\ 0 & -\theta \end{pmatrix}\begin{pmatrix} Y_t \\ X_t \end{pmatrix} + \begin{pmatrix} a \\ m \end{pmatrix}\right]\mathrm{d}t + \begin{pmatrix} \sqrt{Y_t} & 0 \\ 0 & \sqrt{Y_t} \end{pmatrix}\begin{pmatrix} \mathrm{d}L_t \\ \mathrm{d}B_t \end{pmatrix}, \qquad t \geqslant 0.$$

Note also that under the condition $a \geqslant \frac{1}{2}$, for all $y_0 > 0$, we have $\mathbb{P}(Y_t > 0, \forall\, t \in \mathbb{R}_+ \mid Y_0 = y_0) = 1$, see, e.g., page 442 in Revuz and Yor [23]. Since $\mathbb{P}(Y_0 > 0) = 1$, by the law of total probability, $\mathbb{P}(Y_t > 0, \forall\, t \in \mathbb{R}_+) = 1$.

We intend to use formula (7.139) in Section 7.6.4 of Liptser and Shiryaev [19]. Their condition (7.137) has to be checked which takes the form

$$\int_0^T \left(\frac{(a - bY_s)^2}{Y_s} + \frac{(m-\theta X_s)^2}{Y_s}\right)\mathrm{d}s < \infty \qquad \text{a.s. for all } T \in \mathbb{R}_+.$$

Since (Y, X) has continuous sample paths almost surely, this holds if

$$\int_0^T \frac{1}{Y_s}\,\mathrm{d}s < \infty \qquad \text{a.s. for all } T \in \mathbb{R}_+. \tag{3.2}$$

Under the conditions $a \geqslant 1/2$ and $\mathbb{P}(Y_0 > 0) = 1$, Theorems 2 and 4 in Ben-Alaya and Kebaier [5] yield this property. More precisely, if $a \geqslant \frac{1}{2}$ and $y_0 > 0$, then Theorems 2 and 4 in Ben-Alaya and Kebaier [5] yield

$$\mathbb{P}\left(\int_0^T \frac{1}{Y_s}\,\mathrm{d}s < \infty \,\middle|\, Y_0 = y_0\right) = 1, \qquad T \in \mathbb{R}_+.$$

Since $\mathbb{P}(Y_0 > 0) = 1$, by the law of total probability, we get (3.2).

We call the attention that conditions (4.110) and (4.111) are also required for Section 7.6.4 in Liptser and Shiryaev [19], but the Lipschitz condition (4.110) in Liptser and Shiryaev [19] does not hold for the SDE (1.1). However, we can use formula (7.139) in Liptser and Shiryaev [19], since they use their conditions (4.110) and (4.111) only in order to ensure the SDE they consider in Section 7.6.4 has a unique strong solution (see, the proof of Theorem 7.19 in Liptser and Shiryaev [19]).

By Proposition 2.1, under the conditions of the present lemma, there is a (pathwise) unique strong solution of the SDE (1.1).

The martingale property of the process (3.1) is a consequence of Jacod and Shiryaev [15, Chapter III, Theorem 3.4]. $\square$

By Lemma 3.1, under its conditions the log-likelihood function takes the form

$$\begin{aligned}
\log L_T^{(\theta,m),(0,0)}((Y_s,X_s)_{s\in[0,T]}) &= m\int_0^T \frac{1}{Y_s}\,\mathrm{d}X_s - \theta\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s - \frac{m^2}{2}\int_0^T \frac{1}{Y_s}\,\mathrm{d}s \\
&\quad + m\theta\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s - \frac{\theta^2}{2}\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \\
&=: f_T(\theta,m), \qquad T>0.
\end{aligned}$$

We remark that for all $a \geqslant 1/2$, $b\in\mathbb{R}$, $T>0$, and all initial values (Y_0,X_0), the probability measures $\mathbb{P}_{(\theta,m),T}$, $m,\theta\in\mathbb{R}$, are absolutely continuous with respect to each other, and hence it does not matter which measure is taken as a reference measure for defining the maximum likelihood estimator (we have chosen $\mathbb{P}_{(0,0),T}$). For more details, see, e.g., Liptser and Shiryaev [19, page 35]. Then the equation $\frac{\partial f_T}{\partial\theta}(\theta,m)=0$ and the system of equations

$$\frac{\partial f_T}{\partial\theta}(\theta,m)=0, \qquad \frac{\partial f_T}{\partial m}(\theta,m)=0,$$

respectively, take the forms

$$-\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s + m\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s - \theta\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s = 0,$$

and

$$\begin{pmatrix} -\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s & \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s \\ \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s & -\int_0^T \frac{1}{Y_s}\,\mathrm{d}s \end{pmatrix}\begin{pmatrix}\theta\\ m\end{pmatrix} = \begin{pmatrix} \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s \\ -\int_0^T \frac{1}{Y_s}\,\mathrm{d}X_s\end{pmatrix}.$$

First, we suppose that $m\in\mathbb{R}$ is known (besides that $a\geqslant 1/2$ and $b\in\mathbb{R}$ are also supposed to be known). By maximizing $\log L_T^{(\theta,m),(0,0)}$ in $\theta\in\mathbb{R}$, we get the maximum likelihood estimator (MLE) of θ based on the observations $(X_t)_{t\in[0,T]}$,

$$\widetilde{\theta}_T^{\mathrm{MLE}} := \frac{-\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s + m\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s}, \qquad T>0, \tag{3.3}$$

provided that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s>0$. Indeed,

$$\frac{\partial^2 f_T}{\partial\theta^2}(\theta,m) = -\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s < 0.$$

Using the SDE (1.1), one can also get

$$\widetilde{\theta}_T^{\mathrm{MLE}} - \theta = -\frac{\int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s}, \qquad T>0, \tag{3.4}$$

provided that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s > 0$.

Similarly, not supposing that $m \in \mathbb{R}$ is known (but supposing that $a \geqslant 1/2$ and $b \in \mathbb{R}$ are known), by maximizing $\log L_T^{(\theta,m),(0,0)}$ in $(\theta, m) \in \mathbb{R}^2$, the maximum likelihood estimator (MLE) of (θ, m) based on the observations $(X_t)_{t\in[0,T]}$ takes the form

$$\widehat{\theta}_T^{\mathrm{MLE}} := \frac{\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}X_s - \int_0^T \frac{1}{Y_s}\,\mathrm{d}s \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}, \qquad T > 0, \tag{3.5}$$

$$\widehat{m}_T^{\mathrm{MLE}} := \frac{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}X_s - \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}, \qquad T > 0, \tag{3.6}$$

provided that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 > 0$. Indeed,

$$\begin{pmatrix} \frac{\partial^2 f_T}{\partial\theta^2}(\theta,m) & \frac{\partial^2 f_T}{\partial m\partial\theta}(\theta,m) \\ \frac{\partial^2 f_T}{\partial\theta\partial m}(\theta,m) & \frac{\partial^2 f_T}{\partial m^2}(\theta,m) \end{pmatrix} = \begin{pmatrix} -\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s & \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s \\ \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s & -\int_0^T \frac{1}{Y_s}\,\mathrm{d}s \end{pmatrix},$$

and the positivity of $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2$ yields that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s > 0$. Using the SDE (1.1) one can check that

$$\begin{pmatrix} \widehat{\theta}_T^{\mathrm{MLE}} - \theta \\ \widehat{m}_T^{\mathrm{MLE}} - m \end{pmatrix} = \begin{pmatrix} -\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s & \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s \\ \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s & -\int_0^T \frac{1}{Y_s}\,\mathrm{d}s \end{pmatrix}^{-1} \begin{pmatrix} \int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s \\ -\int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s \end{pmatrix},$$

and hence

$$\widehat{\theta}_T^{\mathrm{MLE}} - \theta = \frac{\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s - \int_0^T \frac{1}{Y_s}\,\mathrm{d}s \int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}, \qquad T > 0, \tag{3.7}$$

and

$$\widehat{m}_T^{\mathrm{MLE}} - m = \frac{-\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s \int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s + \int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}, \qquad T > 0, \tag{3.8}$$

provided that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 > 0$.

3.2 Remark. In order that the stochastic integrals $\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s$ and $\int_0^T \frac{1}{Y_s}\,\mathrm{d}X_s$ in (3.3), (3.5) and (3.6) can be observable, we define them pathwise in the following way. With the notations of Jacod and Shiryaev [15], $\left(\tau_n := \left(\frac{i}{n} \wedge T\right)_{i\in\mathbb{N}}\right)_{n\in\mathbb{N}}$ is a Riemann sequence of (adapted) subdivisions and hence, by Jacod and Shiryaev [15, Proposition I.4.44], the sequence of processes

$$\left(\sum_{i=1}^{\lfloor nT\rfloor} \frac{X_{\frac{i-1}{n}\wedge T\wedge t}}{Y_{\frac{i-1}{n}\wedge T\wedge t}} \big(X_{\frac{i}{n}\wedge T\wedge t} - X_{\frac{i-1}{n}\wedge T\wedge t}\big)\right)_{t\in\mathbb{R}_+}, \qquad n \in \mathbb{N},$$

converges to $\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s$ in probability as $n\to\infty$, uniformly on every compact interval. Similarly,

$$\left(\sum_{i=1}^{\lfloor nT\rfloor}\frac{1}{Y_{\frac{i-1}{n}\wedge T\wedge t}}\big(X_{\frac{i}{n}\wedge T\wedge t}-X_{\frac{i-1}{n}\wedge T\wedge t}\big)\right)_{t\in\mathbb{R}_+}, \qquad n\in\mathbb{N},$$

converges to $\int_0^T \frac{1}{Y_s}\,\mathrm{d}X_s$ in probability as $n\to\infty$, uniformly on every compact interval. This means that in practice the continuous time observation $(Y_t,X_t)_{t\in[0,T]}$ is used for giving a discrete time approximation of the appearing stochastic integrals in (3.3), (3.5) and (3.6). $\square$

The next lemma is about the existence of $\widetilde{\theta}_T^{\mathrm{MLE}}$ (supposing that $m\in\mathbb{R}$ is known).

3.3 Lemma. *Let us suppose that $m\in\mathbb{R}$ is known. If $a\geqslant\frac{1}{2}$, $b,\theta\in\mathbb{R}$, and $\mathbb{P}(Y_0>0)=1$, then*

$$\mathbb{P}\left(\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s\in(0,\infty)\right)=1 \quad \text{for all } T>0, \tag{3.9}$$

and hence there exists a unique MLE $\widetilde{\theta}_T^{\mathrm{MLE}}$ which has the form given in (3.3).

Proof. First note that under the condition $a\geqslant\frac{1}{2}$, for all $y_0>0$, we have $\mathbb{P}(Y_t>0,\ \forall\, t\in\mathbb{R}_+\,|\,Y_0=y_0)=1$, see, e.g., page 442 in Revuz and Yor [23]. Since $\mathbb{P}(Y_0>0)=1$, by the law of total probability, we get $\mathbb{P}(Y_t>0,\ \forall\, t\in\mathbb{R}_+)=1$. Note also that, since X has continuous trajectories almost surely, by the proof of Lemma 3.1, we have $\mathbb{P}\Big(\int_0^T\frac{X_s^2}{Y_s}\,\mathrm{d}s\in[0,\infty)\Big)=1$ for all $T>0$. Further, for any $\omega\in\Omega$, $\int_0^T\frac{X_s^2(\omega)}{Y_s(\omega)}\,\mathrm{d}s=0$ holds if and only if $X_s(\omega)=0$ for almost every $s\in[0,T]$. Using that X has continuous trajectories almost surely, we have

$$\mathbb{P}\left(\int_0^T\frac{X_s^2}{Y_s}\,\mathrm{d}s=0\right)>0$$

holds if and only if $\mathbb{P}(X_s=0,\ \forall\, s\in[0,T])>0$. Due to $a\geqslant\frac{1}{2}$ there does not exist a constant $c\in\mathbb{R}$ such that $\mathbb{P}(X_s=c,\ \forall\, s\in[0,T])>0$. Indeed, if $c\in\mathbb{R}$ is such that $\mathbb{P}(X_s=c,\ \forall\, s\in[0,T])>0$, then using (2.2), we have

$$c=\mathrm{e}^{-\theta s}\left(c+m\int_0^s\mathrm{e}^{\theta u}\,\mathrm{d}u+\int_0^s\mathrm{e}^{\theta u}\sqrt{Y_u}\,\mathrm{d}B_u\right), \qquad s\in[0,T],$$

on the event $\{\omega\in\Omega: X_s(\omega)=c,\ \forall\, s\in[0,T]\}$. In case $\theta\neq0$, the process

$$\left(\int_0^s\mathrm{e}^{\theta u}\sqrt{Y_u}\,\mathrm{d}B_u\right)_{s\in[0,T]}$$

would be equal to the deterministic process $\big((c-m/\theta)(\mathrm{e}^{\theta s}-1)\big)_{s\in[0,T]}$ on the event $\{\omega\in\Omega: X_s(\omega)=c,\ \forall\, s\in[0,T]\}$ having positive probability. Since the quadratic variation of the deterministic process $\big((c-m/\theta)(\mathrm{e}^{\theta s}-1)\big)_{s\in[0,T]}$ is the identically zero process (due to the fact that the quadratic variation process is a process starting from 0 almost surely), the quadratic variation of $\big(\int_0^s\mathrm{e}^{\theta u}\sqrt{Y_u}\,\mathrm{d}B_u\big)_{s\in[0,T]}$ should be identically zero on the event $\{\omega\in\Omega: X_s(\omega)=c,\ \forall\, s\in[0,T]\}$. This would imply that $\int_0^s\mathrm{e}^{2\theta u}Y_u\,\mathrm{d}u=0$ for all $s\in[0,T]$ on the event $\{\omega\in\Omega: X_s(\omega)=c,\ \forall\, s\in[0,T]\}$. Using the almost sure continuity and non-negativeness of the sample paths of Y, we have $Y_s=0$ for all $s\in[0,T]$ on the event

$$\{\omega\in\Omega: X_s(\omega)=c,\ \forall\, s\in[0,T]\}\cap\{\omega\in\Omega: (Y_s(\omega))_{s\in[0,T]} \text{ is continuous}\}$$

having positive probability. If $\theta = 0$, then replacing $\big((c - m/\theta)(\mathrm{e}^{\theta s} - 1)\big)_{s\in[0,T]}$ by $(-ms)_{s\in[0,T]}$, one can arrive at the same conclusion. Hence $\mathbb{P}(\inf\{t \in \mathbb{R}_+ : Y_t = 0\} = 0) > 0$ which leads us to a contradiction. This implies (3.9).

The above argument also shows that we can make a shortcut by arguing in a little bit different way. Indeed, if C is a random variable such that $\mathbb{P}(X_s = C, \ \forall\, s \in [0,T]) > 0$, then on the event $\{\omega \in \Omega : X_s(\omega) = C(\omega), \ \forall\, s \in [0,T]\}$, the quadratic variation of X would be identically zero. Since, by the SDE (1.1), the quadratic variation of X is the process $\left(\int_0^t Y_u \,\mathrm{d}u\right)_{t\geqslant 0}$, it would imply that $\int_0^s Y_u \,\mathrm{d}u = 0$ for all $s \in [0,T]$ on the event $\{\omega \in \Omega : X_s(\omega) = C(\omega), \ \forall\, s \in [0,T]\}$. Using the almost sure continuity and non-negativeness of the sample paths of Y, we have $Y_s = 0$ for all $s \in [0,T]$ on the event

$$\{\omega \in \Omega : X_s(\omega) = C(\omega), \ \forall\, s \in [0,T]\} \cap \{\omega \in \Omega : (Y_s(\omega))_{s\in[0,T]} \text{ is continuous}\}$$

having positive probability. As before, this leads us to a contradiction. □

The next lemma is about the existence of $(\widehat{\theta}_T^{\mathrm{MLE}}, \widehat{m}_T^{\mathrm{MLE}})$.

3.4 Lemma. *If* $a \geqslant \frac{1}{2}$, $b, \theta, m \in \mathbb{R}$, *and* $\mathbb{P}(Y_0 > 0) = 1$, *then*

$$\mathbb{P}\left(\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 \in (0,\infty)\right) = 1 \quad \text{for all } T > 0, \tag{3.10}$$

and hence there exists a unique MLE $(\widehat{\theta}_T^{\mathrm{MLE}}, \widehat{m}_T^{\mathrm{MLE}})$ *which has the form given in* (3.5) *and* (3.6).

Proof. First note that under the condition $a \geqslant \frac{1}{2}$, as it was detailed in the proof of Lemma 3.3, we have $\mathbb{P}(Y_t > 0, \ \forall\, t \in \mathbb{R}_+) = 1$. By Cauchy-Schwarz's inequality, we have

$$\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 \geqslant 0, \qquad T > 0, \tag{3.11}$$

and hence, since X has continuous trajectories almost surely, by the proof of Lemma 3.1, we have

$$\mathbb{P}\left(\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 \in [0,\infty)\right) = 1, \qquad T > 0.$$

Further, equality holds in (3.11) if and only if $KX_s^2/Y_s = L/Y_s$ for almost every $s \in [0,T]$ with some $K, L \geqslant 0$, $K^2 + L^2 > 0$ (K and L may depend on $\omega \in \Omega$ and $T > 0$) or equivalently $KX_s^2 = L$ for almost every $s \in [0,T]$ with some $K, L \geqslant 0$, $K^2 + L^2 > 0$. Note that if K were 0, then L would be 0, too, hence K can not be 0 implying that equality holds in (3.11) if and only if $X_s^2 = L/K$ for almost every $s \in [0,T]$ with some $K > 0$ and $L \geqslant 0$ (K and L may depend on $\omega \in \Omega$ and T). Using that X has continuous trajectories almost surely, we have

$$\mathbb{P}\left(\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 = 0\right) > 0 \tag{3.12}$$

holds if and only if $\mathbb{P}(X_s^2 = L/K, \ \forall\, s \in [0,T]) > 0$ with some random variables K and L such that $\mathbb{P}(K > 0, L \geqslant 0) = 1$ (K and L may depend on T). Similarly, as it was explained at the end of the proof of Lemma 3.3, this implies that the quadratic variation of the process $(X_s^2)_{s\in[0,T]}$

would be identically zero on the event $\{\omega \in \Omega : X_s^2(\omega) = L(\omega)/K(\omega), \ \forall\, s \in [0,T]\}$ having positive probability. Since, by Itô's formula,

$$\mathrm{d}X_t^2 = 2X_t\,\mathrm{d}X_t + Y_t\,\mathrm{d}t = (2X_t(m-\theta X_t) + Y_t)\,\mathrm{d}t + 2X_t\sqrt{Y_t}\,\mathrm{d}B_t, \quad t \geqslant 0,$$

the quadratic variation of $(X_t^2)_{t\geqslant 0}$ is the process $\left(\int_0^t 4X_u^2 Y_u\,\mathrm{d}u\right)_{t\geqslant 0}$. If (3.12) holds, then $\int_0^s 4X_u^2 Y_u\,\mathrm{d}u = 0$ for all $s \in [0,T]$ on the event $\{\omega \in \Omega : X_s^2(\omega) = L(\omega)/K(\omega), \ \forall\, s \in [0,T]\}$ having positive probability. Using the almost sure continuity and non-negativeness of the sample paths of $X^2 Y$, we have $X_s^2 Y_s = 0$ for all $s \in [0,T]$ on the event

$$\{\omega \in \Omega : X_s^2(\omega) = L(\omega)/K(\omega), \ \forall\, s \in [0,T]\} \cap \{\omega \in \Omega : (X_t^2(\omega)Y_t(\omega))_{t\geqslant 0} \text{ is continuous}\} =: A$$

having positive probability. Since $\mathbb{P}(Y_t > 0, \ \forall\, t \geqslant 0) = 1$, we have $X_s = 0$ for all $s \in [0,T]$ on the event A having positive probability. Repeating the argument given in the proof of Lemma 3.3, we arrive at a contradiction. This implies (3.10). □

4 Existence and uniqueness of least squares estimator

First we give a motivation for the least squares estimator (LSE) based continuous time observations using the form of the LSE based on discrete time observations.

Let us suppose that $m \in \mathbb{R}$ is known (besides that $a > 0$ and $b \in \mathbb{R}$ are also supposed to be known). The LSE of θ based on the discrete time observations X_i, $i = 0, 1, \ldots, n$, can be obtained by solving the following extremum problem

$$\widetilde{\theta}_n^{\mathrm{LSE,D}} := \operatorname*{arg\,min}_{\theta\in\mathbb{R}} \sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1}))^2.$$

Here in the notation $\widetilde{\theta}_n^{\mathrm{LSE,D}}$ the letter D refers to discrete time observations, and we note that X_0 denotes an observation for the second coordinate of the initial value of the process (Y, X). This definition of LSE of θ can be considered as the corresponding one given in Hu and Long [13, formula (1.2)] for generalized Ornstein-Uhlenbeck processes driven by α-stable motions, see also Hu and Long [14, formula (3.1)]. For a mathematical motivation of the LSE of θ based on the discrete observations X_i, $i = 0, 1, \ldots, n$, see Remark 3.4 in Barczy et al. [1]. Further, by Barczy et al. [1, formula (3.5)],

$$\widetilde{\theta}_n^{\mathrm{LSE,D}} = -\frac{\sum_{i=1}^n (X_i - X_{i-1})X_{i-1} - m\left(\sum_{i=1}^n X_{i-1}\right)}{\sum_{i=1}^n X_{i-1}^2} \tag{4.1}$$

provided that $\sum_{i=1}^n X_{i-1}^2 > 0$. Motivated by (4.1), the LSE of θ based on the continuous time observations $(X_t)_{t\in[0,T]}$ is defined by

$$\widetilde{\theta}_T^{\mathrm{LSE}} := -\frac{\int_0^T X_s\,\mathrm{d}X_s - m\int_0^T X_s\,\mathrm{d}s}{\int_0^T X_s^2\,\mathrm{d}s}, \tag{4.2}$$

provided that $\int_0^T X_s^2\,\mathrm{d}s > 0$, and using the SDE (1.1) we have

$$\widetilde{\theta}_T^{\mathrm{LSE}} - \theta = -\frac{\int_0^T X_s\sqrt{Y_s}\,\mathrm{d}B_s}{\int_0^T X_s^2\,\mathrm{d}s}, \tag{4.3}$$

provided that $\int_0^T X_s^2 \, \mathrm{d}s > 0$.

Let us suppose that $m \in \mathbb{R}$ is not known (but supposing that $a > 0$ and $b \in \mathbb{R}$ are known). The LSE of (θ, m) based on the discrete time observations X_i, $i = 0, 1, \ldots, n$, can be obtained by solving the following extremum problem

$$(\widehat{\theta}_n^{\mathrm{LSE,D}}, \widehat{m}_n^{\mathrm{LSE,D}}) := \operatorname*{arg\,min}_{(\theta,m)\in\mathbb{R}^2} \sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1}))^2.$$

By Barczy et al. [1, formulas (3.16) and (3.17)],

$$\widehat{\theta}_n^{\mathrm{LSE,D}} = \frac{\sum_{i=1}^n X_{i-1} \sum_{i=1}^n (X_i - X_{i-1}) - n \sum_{i=1}^n (X_i - X_{i-1}) X_{i-1}}{n \sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2}, \tag{4.4}$$

and

$$\widehat{m}_n^{\mathrm{LSE,D}} = \frac{\sum_{i=1}^n X_{i-1}^2 \sum_{i=1}^n (X_i - X_{i-1}) - \sum_{i=1}^n X_{i-1} \sum_{i=1}^n (X_i - X_{i-1}) X_{i-1}}{n \sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2} \tag{4.5}$$

provided that $n \sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2 > 0$. Motivated by (4.4) and (4.5), the LSE of (θ, m) based on the continuous time observations $(X_t)_{t\in[0,T]}$ is defined by

$$\widehat{\theta}_T^{\mathrm{LSE}} := \frac{(X_T - X_0) \int_0^T X_s \, \mathrm{d}s - T \int_0^T X_s \, \mathrm{d}X_s}{T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s\right)^2}, \tag{4.6}$$

$$\widehat{m}_T^{\mathrm{LSE}} := \frac{(X_T - X_0) \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s\right) \left(\int_0^T X_s \, \mathrm{d}X_s\right)}{T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s\right)^2}, \tag{4.7}$$

provided that $T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s\right)^2 > 0$. Note that, by Cauchy-Schwarz's inequality, $T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s\right)^2 \geqslant 0$, and $T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s\right)^2 > 0$ yields that $\int_0^T X_s^2 \, \mathrm{d}s > 0$. Then

$$\begin{pmatrix} \widehat{\theta}_T^{\mathrm{LSE}} \\ \widehat{m}_T^{\mathrm{LSE}} \end{pmatrix} = \begin{pmatrix} -\int_0^T X_s^2 \, \mathrm{d}s & \int_0^T X_s \, \mathrm{d}s \\ \int_0^T X_s \, \mathrm{d}s & -T \end{pmatrix}^{-1} \begin{pmatrix} \int_0^T X_s \, \mathrm{d}X_s \\ -(X_T - X_0) \end{pmatrix},$$

and using the SDE (1.1) one can check that

$$\begin{pmatrix} \widehat{\theta}_T^{\mathrm{LSE}} - \theta \\ \widehat{m}_T^{\mathrm{LSE}} - m \end{pmatrix} = \begin{pmatrix} -\int_0^T X_s^2 \, \mathrm{d}s & \int_0^T X_s \, \mathrm{d}s \\ \int_0^T X_s \, \mathrm{d}s & -T \end{pmatrix}^{-1} \begin{pmatrix} \int_0^T X_s \sqrt{Y_s} \, \mathrm{d}B_s \\ -\int_0^T \sqrt{Y_s} \, \mathrm{d}B_s \end{pmatrix},$$

and hence

$$\widehat{\theta}_T^{\mathrm{LSE}} - \theta = \frac{-T \int_0^T X_s \sqrt{Y_s} \, \mathrm{d}B_s + \int_0^T X_s \, \mathrm{d}s \int_0^T \sqrt{Y_s} \, \mathrm{d}B_s}{T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s\right)^2}, \tag{4.8}$$

and

$$\widehat{m}_T^{\mathrm{LSE}} - m = \frac{-\int_0^T X_s \, \mathrm{d}s \int_0^T X_s \sqrt{Y_s} \, \mathrm{d}B_s + \int_0^T X_s^2 \, \mathrm{d}s \int_0^T \sqrt{Y_s} \, \mathrm{d}B_s}{T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s\right)^2}, \tag{4.9}$$

provided that $T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s\right)^2 > 0$.

4.1 Remark. In order that the stochastic integral $\int_0^T X_s \, \mathrm{d}X_s$ in (4.2), (4.6) and (4.7) can be observable, we define it pathwise in the following way. By Itô's formula, we have $\mathrm{d}X_t^2 = 2X_t \, \mathrm{d}X_t + Y_t \, \mathrm{d}t$, $t \in \mathbb{R}_+$, and then

$$\int_0^T X_s \, \mathrm{d}X_s = \frac{1}{2} \left(X_T^2 - X_0^2 - \int_0^T Y_s \, \mathrm{d}s \right).$$

□

The next lemma is about the existence of $\widetilde{\theta}_T^{\mathrm{LSE}}$ (supposing that $m \in \mathbb{R}$ is known).

4.2 Lemma. *Let us suppose that $m \in \mathbb{R}$ is known. If $b, \theta \in \mathbb{R}$ and $a > 0$, then*

$$\mathbb{P}\left(\int_0^T X_s^2 \, \mathrm{d}s \in (0, \infty) \right) = 1 \quad \text{for all } T > 0, \tag{4.10}$$

and hence there exists a unique LSE $\widetilde{\theta}_T^{\mathrm{LSE}}$ which has the form given in (4.2).

Proof. First note that under the condition on the parameters, for all $y_0 > 0$, we have

$$\mathbb{P}(\inf\{t \in \mathbb{R}_+ : Y_t = 0\} > 0 \,|\, Y_0 = y_0) = 1,$$

and hence, by the law of total probability, $\mathbb{P}(\inf\{t \in \mathbb{R}_+ : Y_t = 0\} > 0) = 1$. Since X has continuous trajectories almost surely, we readily have

$$\mathbb{P}\left(\int_0^T X_s^2 \, \mathrm{d}s \in [0, \infty) \right) = 1, \qquad T > 0.$$

Observe also that for any $\omega \in \Omega$, $\int_0^T X_s^2(\omega) \, \mathrm{d}s = 0$ holds if and only if $X_s(\omega) = 0$ for almost every $s \in [0, T]$. Using that X has continuous trajectories almost surely, we have

$$\mathbb{P}\left(\int_0^T X_s^2 \, \mathrm{d}s = 0 \right) > 0$$

holds if and only if $\mathbb{P}(X_s = 0, \, \forall \, s \in [0, T]) > 0$. By the end of the proof of Lemma 3.3, if $\mathbb{P}(X_s = 0, \, \forall \, s \in [0, T]) > 0$, then $Y_s = 0$ for all $s \in [0, T]$ on the event

$$\Big\{ \omega \in \Omega : X_s(\omega) = 0, \, \forall \, s \in [0, T] \Big\} \cap \big\{ \omega \in \Omega : (Y_s(\omega))_{s \in [0,T]} \text{ is continuous} \big\}$$

having positive probability. This would yield that $\mathbb{P}(\inf\{t \in \mathbb{R}_+ : Y_t = 0\} = 0) > 0$ leading us to a contradiction, which implies (4.10). □

The next lemma is about the existence of $(\widehat{\theta}_T^{\mathrm{LSE}}, \widehat{m}_T^{\mathrm{LSE}})$.

4.3 Lemma. *If $b, \theta \in \mathbb{R}$ and $a > 0$, then*

$$\mathbb{P}\left(T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s \right)^2 \in (0, \infty) \right) = 1 \quad \text{for all } T > 0, \tag{4.11}$$

and hence there exists a unique LSE $(\widehat{\theta}_T^{\mathrm{LSE}}, \widehat{m}_T^{\mathrm{LSE}})$ which has the form given in (4.6) *and* (4.7).

Proof. Just as in the proof of Lemma 4.2, we have $\mathbb{P}(\inf\{t \in \mathbb{R}_+ : Y_t = 0\} > 0) = 1$. By Cauchy-Schwarz's inequality, we have

$$T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s \right)^2 \geqslant 0, \qquad T > 0, \tag{4.12}$$

and hence, since X has continuous trajectories almost surely, we readily have

$$\mathbb{P}\left(T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s \right)^2 \in [0, \infty) \right) = 1, \qquad T > 0.$$

Further, equality holds in (4.12) if and only if $KX_s^2 = L$ for almost every $s \in [0, T]$ with some $K, L \geqslant 0$, $K^2 + L^2 > 0$ (K and L may depend on $\omega \in \Omega$ and T). Note that if K were 0, then L would be 0, too, hence K can not be 0 implying that equality holds in (4.12) if and only if $X_s^2 = L/K$ for almost every $s \in [0, T]$ with some $K > 0$ and $L \geqslant 0$ (K and L may depend on $\omega \in \Omega$ and T). Using that X has continuous trajectories almost surely, we have

$$\mathbb{P}\left(T \int_0^T X_s^2 \, \mathrm{d}s - \left(\int_0^T X_s \, \mathrm{d}s \right)^2 = 0 \right) > 0 \tag{4.13}$$

holds if and only if $\mathbb{P}(X_s^2 = L/K, \ \forall \, s \in [0, T]) > 0$ with some random variables K and L such that $\mathbb{P}(K > 0, L \geqslant 0) = 1$ (K and L may depend on T). Using again that X has continuous trajectories almost surely, we have (4.13) holds if and only if $\mathbb{P}(X_s = C, \ \forall \, s \in [0, T]) > 0$ with some random variable C (note that $C = \sqrt{L/K}$ if X is non-negative and $C = -\sqrt{L/K}$ if X is negative). This leads us to a contradiction by the end of the proof of Lemma 3.3. Indeed, if $\mathbb{P}(X_s = C, \ \forall \, s \in [0, T]) > 0$ with some random variable C, then, by the end of the proof of Lemma 3.3, we have $Y_s = 0$ for all $s \in [0, T]$ on the event

$$\Big\{\omega \in \Omega : X_s(\omega) = C(\omega), \ \forall \, s \in [0, T]\Big\} \cap \big\{\omega \in \Omega : (Y_s(\omega))_{s \in [0,T]} \ \text{ is continuous}\big\}$$

having positive probability. This would yield that $\mathbb{P}(\inf\{t \in \mathbb{R}_+ : Y_t = 0\} = 0) > 0$ leading us to a contradiction, which implies (4.11). $\square$

5 Consistency of maximum likelihood estimator

5.1 Theorem. *Let us suppose that $m \in \mathbb{R}$ is known. If $a \geqslant \frac{1}{2}$, $b > 0$, $\theta > 0$, and $\mathbb{P}(Y_0 > 0) = 1$, then the MLE of θ is strongly consistent:* $\mathbb{P}\left(\lim_{T \to \infty} \widetilde{\theta}_T^{\mathrm{MLE}} = \theta \right) = 1$.

Proof. By Lemma 3.3, there exists a unique MLE $\widetilde{\theta}_T^{\mathrm{MLE}}$ of θ which has the form given in (3.3). Next we prove that

$$\mathbb{P}\left(\lim_{T \to \infty} \int_0^T \frac{X_s^2}{Y_s} \, \mathrm{d}s = +\infty \right) = 1. \tag{5.1}$$

By Theorem 2.3, we have

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \frac{X_s^2}{1 + Y_s} \, \mathrm{d}s = \mathbb{E}\left(\frac{X_\infty^2}{1 + Y_\infty} \right) \qquad \text{a.s.},$$

where

$$\mathbb{E}\left(\frac{X_\infty^2}{1+Y_\infty}\right) \leqslant \mathbb{E}(X_\infty^2) = \frac{a\theta + 2bm^2}{2b\theta^2} < \infty.$$

Next we check that $\mathbb{E}(X_\infty^2/(1+Y_\infty)) > 0$. Since $X_\infty^2/(1+Y_\infty) \geqslant 0$, we have $\mathbb{E}(X_\infty^2/(1+Y_\infty)) = 0$ holds if and only if $\mathbb{P}(X_\infty^2/(1+Y_\infty) = 0) = 1$ or equivalently $\mathbb{P}(X_\infty = 0) = 1$ which leads us to a contradiction since X_∞ is absolutely continuous by Theorem 2.3. Hence we have

$$\mathbb{P}\left(\lim_{T\to\infty} \int_0^T \frac{X_s^2}{1+Y_s}\,\mathrm{d}s = \infty\right) = 1,$$

which yields (5.1). Then, using (3.4), a strong law of large numbers for continuous local martingales (see, e.g., Theorem 2.5) yields the strong consistency of $\widetilde{\theta}_T^{\mathrm{MLE}}$. □

5.2 Theorem. *If* $a > \frac{1}{2}$, $b > 0$, $\theta > 0$, $m \in \mathbb{R}$, *and* $\mathbb{P}(Y_0 > 0) = 1$, *then the MLE of* (θ, m) *is strongly consistent:* $\mathbb{P}\left(\lim_{T\to\infty}(\widehat{\theta}_T^{\mathrm{MLE}}, \widehat{m}_T^{\mathrm{MLE}}) = (\theta, m)\right) = 1$.

Proof. By Lemma 3.4, there exists a unique MLE $(\widehat{\theta}_T^{\mathrm{MLE}}, \widehat{m}_T^{\mathrm{MLE}})$ of (θ, m) which has the form given in (3.5) and (3.6). By (3.7), using also that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 > 0$ yields $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s > 0$, we have

$$\widehat{\theta}_T^{\mathrm{MLE}} - \theta = \frac{\frac{\frac{1}{T}\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s}{\frac{1}{T}\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s} \cdot \frac{\int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\int_0^T \frac{1}{Y_s}\,\mathrm{d}s} - \frac{\int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s}}{1 - \frac{\left(\frac{1}{T}\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}{\frac{1}{T}\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \cdot \frac{1}{T}\int_0^T \frac{1}{Y_s}\,\mathrm{d}s}} \tag{5.2}$$

provided that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 > 0$. Next, we show that $\mathbb{E}(1/Y_\infty) < \infty$, $\mathbb{E}(X_\infty/Y_\infty) < \infty$ and $\mathbb{E}(X_\infty^2/Y_\infty) < \infty$, which, by part (iii) of Theorem 2.3, will imply that

$$\mathbb{P}\left(\lim_{T\to\infty} \frac{1}{T}\int_0^T \frac{1}{Y_s}\,\mathrm{d}s = \mathbb{E}\left(\frac{1}{Y_\infty}\right)\right) = 1, \tag{5.3}$$

$$\mathbb{P}\left(\lim_{T\to\infty} \frac{1}{T}\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s = \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)\right) = 1, \tag{5.4}$$

$$\mathbb{P}\left(\lim_{T\to\infty} \frac{1}{T}\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s = \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)\right) = 1. \tag{5.5}$$

We only prove $\mathbb{E}(X_\infty^2/Y_\infty) < \infty$ noting that $\mathbb{E}(1/Y_\infty) < \infty$ and $\mathbb{E}(X_\infty/Y_\infty) < \infty$ can be checked in the very same way. First note that, by Theorem 2.3, $\mathbb{P}(Y_\infty > 0) = 1$, and hence the random variable X_∞^2/Y_∞ is well-defined with probability one. By Hölder's inequality, for all $p, q > 0$ with $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) \leqslant (\mathbb{E}(X_\infty^{2p}))^{1/p}\left(\mathbb{E}\left(\frac{1}{Y_\infty^q}\right)\right)^{1/q}.$$

Since, by Theorem 2.3, $\mathbb{E}(|X_\infty|^n) < \infty$ for all $n \in \mathbb{N}$, to prove that $\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) < \infty$ it is enough to check that there exists some $\varepsilon > 0$ such that $\mathbb{E}\left(\frac{1}{Y_\infty^{1+\varepsilon}}\right) < \infty$. Using that, by Theorem 2.3, Y_∞

has Gamma distribution with density function $\frac{(2b)^{2a}}{\Gamma(2a)}x^{2a-1}\mathrm{e}^{-2bx}\mathbf{1}_{\{x>0\}}$, we have, for any $\varepsilon > 0$,

$$\mathbb{E}\left(\frac{1}{Y_\infty^{1+\varepsilon}}\right) = \int_0^\infty \frac{1}{x^{1+\varepsilon}} \cdot \frac{(2b)^{2a}}{\Gamma(2a)} x^{2a-1}\mathrm{e}^{-2bx}\,\mathrm{d}x = \frac{(2b)^{2a}}{\Gamma(2a)} \int_0^\infty x^{2a-2-\varepsilon}\mathrm{e}^{-2bx}\,\mathrm{d}x.$$

Due to our assumption $a > 1/2$, one can choose an ε such that $\max(0, 2a-2) < \varepsilon < 2a-1$, and hence for all $M > 0$,

$$\begin{aligned}
\int_0^\infty x^{2a-2-\varepsilon}\mathrm{e}^{-2bx}\,\mathrm{d}x &\leqslant \int_0^M x^{2a-2-\varepsilon}\,\mathrm{d}x + M^{2a-2-\varepsilon}\int_M^\infty \mathrm{e}^{-2bx}\,\mathrm{d}x \\
&= \frac{M^{2a-1-\varepsilon}}{2a-1-\varepsilon} + M^{2a-2-\varepsilon} \lim_{L\to\infty} \frac{\mathrm{e}^{-2bL} - \mathrm{e}^{-2bM}}{-2b} \\
&= \frac{M^{2a-1-\varepsilon}}{2a-1-\varepsilon} + M^{2a-2-\varepsilon}\frac{\mathrm{e}^{-2bM}}{2b} < \infty,
\end{aligned}$$

which yields that $\mathbb{E}(1/Y_\infty^{1+\varepsilon}) < \infty$ for ε satisfying $\max(0, 2a-2) < \varepsilon < 2a-1$.

Further, since $\mathbb{E}(1/Y_\infty) > 0$ and $\mathbb{E}(X_\infty^2/Y_\infty) > 0$ (due to the absolutely continuity of X_∞, as it was explained in the proof of Theorem 5.1), (5.3) and (5.5) yield that

$$\mathbb{P}\left(\lim_{T\to\infty}\int_0^T \frac{1}{Y_s}\,\mathrm{d}s = \infty\right) = \mathbb{P}\left(\lim_{T\to\infty}\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s = \infty\right) = 1.$$

Using (5.2), (5.3), (5.4), (5.5), Slutsky's lemma and a strong law of large numbers for continuous local martingales (see, e.g., Theorem 2.5), we get

$$\lim_{T\to\infty}\left(\widehat{\theta}_T^{\mathrm{MLE}} - \theta\right) = \frac{\frac{\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)}{\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)}\cdot 0 - 0}{1 - \frac{\left(\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)\right)^2}{\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)\mathbb{E}\left(\frac{1}{Y_\infty}\right)}} = 0 \qquad \text{a.s.,}$$

where we also used that the denominator is strictly positive. Indeed, by Cauchy-Schwarz's inequality,

$$\left(\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)\right)^2 \leqslant \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)\mathbb{E}\left(\frac{1}{Y_\infty}\right), \tag{5.6}$$

and equality would hold if and only if $\mathbb{P}(KX_\infty^2/Y_\infty = L/Y_\infty) = 1$ with some $K, L \geqslant 0$, $K^2 + L^2 > 0$. By Theorem 2.3, (Y_∞, X_∞) is absolutely continuous and then $\mathbb{P}(X_\infty = c) = 0$ for all $c \in \mathbb{R}$. This implies that (5.6) holds with strict inequality.

Similarly, by (3.8), we have

$$\widehat{m}_T^{\mathrm{MLE}} - m = \frac{-\frac{\frac{1}{T}\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s}{\frac{1}{T}\int_0^T \frac{1}{Y_s}\,\mathrm{d}s}\cdot\frac{\int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s} + \frac{\int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\int_0^T \frac{1}{Y_s}\,\mathrm{d}s}}{1 - \frac{\left(\frac{1}{T}\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}{\frac{1}{T}\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s\cdot\frac{1}{T}\int_0^T \frac{1}{Y_s}\,\mathrm{d}s}}$$

provided that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 > 0$. Similarly as before one can argue that

$$\lim_{T\to\infty}(\widehat{m}_T^{\mathrm{MLE}} - m) = \frac{-\frac{\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)}{\mathbb{E}\left(\frac{1}{Y_\infty}\right)}\cdot 0 + 0}{1 - \frac{\left(\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)\right)^2}{\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)\mathbb{E}\left(\frac{1}{Y_\infty}\right)}} = 0 \qquad \text{a.s.}$$

Using that the intersection of two events with probability one is an event with probability one, we have the assertion. □

5.3 Remark. If $a = \frac{1}{2}$, $b > 0$, $\theta > 0$, $m \in \mathbb{R}$, and $\mathbb{P}(Y_0 > 0) = 1$, then one should find another approach for studying the consistency behaviour of the MLE of (θ, m), since in this case

$$\mathbb{E}\left(\frac{1}{Y_\infty}\right) = \int_0^\infty \frac{2b\mathrm{e}^{-2bx}}{x}\,\mathrm{d}x = \infty,$$

and hence one cannot use part (iii) of Theorem 2.3. In this paper we renounce to consider it. □

6 Consistency of least squares estimator

6.1 Theorem. *Let us suppose that $m \in \mathbb{R}$ is known. If $a > 0$, $b > 0$, and $\theta > 0$, then the LSE of θ is strongly consistent:* $\mathbb{P}\left(\lim_{T\to\infty} \widetilde{\theta}_T^{\mathrm{LSE}} = \theta\right) = 1$.

Proof. By Lemma 4.2, there exists a unique $\widetilde{\theta}_T^{\mathrm{LSE}}$ of θ which has the form given in (4.2). By (4.3), we have

$$\widetilde{\theta}_T^{\mathrm{LSE}} - \theta = -\frac{\int_0^T X_s\sqrt{Y_s}\,\mathrm{d}B_s}{\int_0^T X_s^2\,\mathrm{d}s} = -\frac{\int_0^T X_s\sqrt{Y_s}\,\mathrm{d}B_s}{\int_0^T X_s^2 Y_s\,\mathrm{d}s}\cdot\frac{\frac{1}{T}\int_0^T X_s^2 Y_s\,\mathrm{d}s}{\frac{1}{T}\int_0^T X_s^2\,\mathrm{d}s}. \tag{6.1}$$

By Theorem 2.3, we have

$$\mathbb{P}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T X_s^2 Y_s\,\mathrm{d}s = \mathbb{E}\left(X_\infty^2 Y_\infty\right)\right) = 1,$$

$$\mathbb{P}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T X_s^2\,\mathrm{d}s = \mathbb{E}\left(X_\infty^2\right)\right) = 1.$$

We note that $\mathbb{E}(X_\infty^2 Y_\infty)$ and $\mathbb{E}(X_\infty^2)$ are calculated explicitly in Theorem 2.3. Note also that $\mathbb{E}(X_\infty^2 Y_\infty)$ is positive (due to that $X_\infty^2 Y_\infty$ is non-negative and absolutely continuous), and hence we also have

$$\mathbb{P}\left(\lim_{T\to\infty}\int_0^T X_s^2 Y_s\,\mathrm{d}s = \infty\right) = 1.$$

Then, by a strong law of large numbers for continuous local martingales (see, e.g., Theorem 2.5), we get

$$\mathbb{P}\left(\lim_{T\to\infty}\frac{\int_0^T X_s\sqrt{Y_s}\,\mathrm{d}B_s}{\int_0^T X_s^2 Y_s\,\mathrm{d}s} = 0\right) = 1,$$

and hence (6.1) yields the assertion. □

6.2 Theorem. *If* $a > 0, b > 0,\ \theta > 0,$ *and* $m \in \mathbb{R}$, *then the LSE of* (θ, m) *is strongly consistent:* $\mathbb{P}\left(\lim_{T\to\infty}(\widehat{\theta}_T^{\mathrm{LSE}}, \widehat{m}_T^{\mathrm{LSE}}) = (\theta, m)\right) = 1.$

Proof. By Lemma 4.3, there exists a unique LSE $(\widehat{\theta}_T^{\mathrm{LSE}}, \widehat{m}_T^{\mathrm{LSE}})$ of (θ, m) which has the form given in (4.6) and (4.7). By (4.8), we have

$$\widehat{\theta}_T^{\mathrm{LSE}} - \theta = \frac{-\frac{\frac{1}{T}\int_0^T X_s^2 Y_s \,\mathrm{d}s}{\frac{1}{T}\int_0^T X_s^2 \,\mathrm{d}s} \cdot \frac{\int_0^T X_s \sqrt{Y_s}\,\mathrm{d}B_s}{\int_0^T X_s^2 Y_s \,\mathrm{d}s} + \frac{1}{T}\int_0^T X_s \,\mathrm{d}s \cdot \frac{\frac{1}{T}\int_0^T Y_s \,\mathrm{d}s}{\frac{1}{T}\int_0^T X_s^2 \,\mathrm{d}s} \cdot \frac{\int_0^T \sqrt{Y_s}\,\mathrm{d}B_s}{\int_0^T Y_s \,\mathrm{d}s}}{1 - \frac{\left(\frac{1}{T}\int_0^T X_s \,\mathrm{d}s\right)^2}{\frac{1}{T}\int_0^T X_s^2 \,\mathrm{d}s}}$$

provided that $T\int_0^T X_s^2 \,\mathrm{d}s - \left(\int_0^T X_s \,\mathrm{d}s\right)^2 > 0$. Since, by Theorem 2.3, $\mathbb{E}(Y_\infty) < \infty$, $\mathbb{E}(X_\infty) < \infty$, $\mathbb{E}(X_\infty^2) < \infty$, and $\mathbb{E}(X_\infty^2 Y_\infty) < \infty$, part (iii) of Theorem 2.3 yields that

$$\begin{aligned}
&\mathbb{P}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T Y_s \,\mathrm{d}s = \mathbb{E}(Y_\infty)\right) = \mathbb{P}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T X_s \,\mathrm{d}s = \mathbb{E}(X_\infty)\right) \\
&= \mathbb{P}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T X_s^2 \,\mathrm{d}s = \mathbb{E}(X_\infty^2)\right) = \mathbb{P}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T X_s^2 Y_s \,\mathrm{d}s = \mathbb{E}(X_\infty^2 Y_\infty)\right) = 1.
\end{aligned}$$

Hence, similarly to the proof of Theorem 5.2, we get

$$\lim_{T\to\infty}(\widehat{\theta}_T^{\mathrm{LSE}} - \theta) = \frac{-\frac{\mathbb{E}(X_\infty^2 Y_\infty)}{\mathbb{E}(X_\infty^2)} \cdot 0 + \mathbb{E}(X_\infty)\cdot\frac{\mathbb{E}(Y_\infty)}{\mathbb{E}(X_\infty^2)}\cdot 0}{1 - \frac{(\mathbb{E}(X_\infty))^2}{\mathbb{E}(X_\infty^2)}} = 0 \qquad \text{a.s.},$$

where for the last step we also used that $(\mathbb{E}(X_\infty))^2 < \mathbb{E}(X_\infty^2)$ (which holds since there does not exist a constant $c \in \mathbb{R}$ such that $\mathbb{P}(X_\infty = c) = 1$ due to the fact that X_∞ is absolutely continuous).

Similarly, by (4.9),

$$\begin{aligned}
\lim_{T\to\infty}(\widehat{m}_T^{\mathrm{LSE}} - m) &= \lim_{T\to\infty}\frac{-\frac{1}{T}\int_0^T X_s \,\mathrm{d}s \cdot \frac{\frac{1}{T}\int_0^T X_s^2 Y_s \,\mathrm{d}s}{\frac{1}{T}\int_0^T X_s^2 \,\mathrm{d}s} \cdot \frac{\int_0^T X_s\sqrt{Y_s}\,\mathrm{d}B_s}{\int_0^T X_s^2 Y_s \,\mathrm{d}s} + \frac{1}{T}\int_0^T Y_s \,\mathrm{d}s \cdot \frac{\int_0^T \sqrt{Y_s}\,\mathrm{d}B_s}{\int_0^T Y_s \,\mathrm{d}s}}{1 - \frac{\left(\frac{1}{T}\int_0^T X_s \,\mathrm{d}s\right)^2}{\frac{1}{T}\int_0^T X_s^2 \,\mathrm{d}s}} \\
&= \frac{-\mathbb{E}(X_\infty)\cdot\frac{\mathbb{E}(X_\infty^2 Y_\infty)}{\mathbb{E}(X_\infty^2)}\cdot 0 + \mathbb{E}(Y_\infty)\cdot 0}{1 - \frac{(\mathbb{E}(X_\infty))^2}{\mathbb{E}(X_\infty^2)}} = 0 \qquad \text{a.s.}
\end{aligned}$$

Using that the intersection of two events with probability one is an event with probability one, we have the assertion. □

7 Asymptotic behaviour of maximum likelihood estimator

7.1 Theorem. *Let us suppose that* $m \in \mathbb{R}$ *is known. If* $a > 1/2$, $b > 0$, $\theta > 0$, *and* $\mathbb{P}(Y_0 > 0) = 1$, *then the MLE of* θ *is asymptotically normal, i.e.,*

$$\sqrt{T}(\widetilde{\theta}_T^{\mathrm{MLE}} - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)}\right) \qquad \text{as } T \to \infty,$$

where $\mathbb{E}\left(X_\infty^2/Y_\infty\right)$ *is positive and finite.*

Proof. First note that, by (3.4),

$$\sqrt{T}(\widetilde{\theta}_T^{\mathrm{MLE}} - \theta) = -\frac{\frac{1}{\sqrt{T}}\int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\frac{1}{T}\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s}, \qquad T > 0, \tag{7.1}$$

provided that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s > 0$. Recall that in the proof of Theorem 5.2 it was shown that (under the conditions of the theorem) $\mathbb{E}\left(X_\infty^2/Y_\infty\right) < \infty$, and hence, by Theorem 2.3, (5.5) holds. Further, since $\mathbb{E}\left(X_\infty^2/Y_\infty\right) > 0$ (due to that X_∞^2/Y_∞ is non-negative and absolutely continuous), (5.5) yields that

$$\mathbb{P}\left(\lim_{T\to\infty}\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s = +\infty\right) = 1.$$

Using Theorem 2.6 with the following choices

$$d := 1, \qquad M_t := \int_0^t \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s, \quad t \geqslant 0, \qquad (\mathcal{F}_t)_{t\geqslant 0} \text{ defined before Proposition 2.1,}$$

$$Q(t) := 1/\sqrt{t}, \quad t > 0, \qquad \eta := \sqrt{\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)},$$

we have

$$\frac{1}{\sqrt{T}}\int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s \xrightarrow{\mathcal{L}} \sqrt{\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)}\,\xi \qquad \text{as } T\to\infty,$$

where ξ is a standard normally distributed random variable. Hence Slutsky's lemma, (5.5) and (7.1) yield the assertion. □

7.2 Theorem. *If $a > 1/2$, $b > 0$, $\theta > 0$, $m \in \mathbb{R}$, and $\mathbb{P}(Y_0 > 0) = 1$, then the MLE of (θ, m) is asymptotically normal, i.e.,*

$$\sqrt{T}\begin{pmatrix}\widehat{\theta}_T^{\mathrm{MLE}} - \theta \\ \widehat{m}_T^{\mathrm{MLE}} - m\end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}_2\left(0, D^{\mathrm{MLE}}\right) \qquad \textit{as } T\to\infty,$$

where $\mathcal{N}_2\left(0, D^{\mathrm{MLE}}\right)$ denotes a 2-dimensional normally distribution with mean vector $0 \in \mathbb{R}^2$ and with covariance matrix

$$D^{\mathrm{MLE}} := \frac{1}{\mathbb{E}\left(\frac{1}{Y_\infty}\right)\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) - \left(\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)\right)^2}\begin{pmatrix}\mathbb{E}\left(\frac{1}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)\end{pmatrix}.$$

Proof. By (3.7) and (3.8), we have

$$\sqrt{T}(\widehat{\theta}_T^{\mathrm{MLE}} - \theta) = \frac{\frac{1}{T}\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\frac{1}{\sqrt{T}}\int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s - \frac{1}{T}\int_0^T \frac{1}{Y_s}\,\mathrm{d}s\frac{1}{\sqrt{T}}\int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\frac{1}{T}\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s\frac{1}{T}\int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\frac{1}{T}\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}, \qquad T > 0,$$

and

$$\sqrt{T}(\widehat{m}_T^{\mathrm{MLE}} - m) = \frac{-\frac{1}{T}\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\frac{1}{\sqrt{T}}\int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s + \frac{1}{T}\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s\frac{1}{\sqrt{T}}\int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\frac{1}{T}\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s\frac{1}{T}\int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\frac{1}{T}\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}, \qquad T > 0,$$

provided that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 > 0$. Next, we show that

$$\left(\frac{1}{\sqrt{T}}\int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s, \frac{1}{\sqrt{T}}\int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s\right) \xrightarrow{\mathcal{L}} \big((\eta Z)_1, (\eta Z)_2\big) \quad \text{as} \quad T \to \infty, \tag{7.2}$$

where Z is a 2-dimensional standard normally distributed random variable and η is a non-random 2×2 matrix such that

$$\eta\eta^\top = \begin{pmatrix} \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{1}{Y_\infty}\right) \end{pmatrix}.$$

Here the matrix

$$\begin{pmatrix} \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{1}{Y_\infty}\right) \end{pmatrix}$$

is positive definite, since its principal minors are positive, and η denotes its unique positive definite square root. Indeed, by the absolutely continuity of (Y_∞, X_∞) (see Theorem 2.3), there does not exist a constant $c \in \mathbb{R}_+$ such that $\mathbb{P}(X_\infty^2 = c) = 1$ and, by Cauchy-Schwarz's inequality,

$$\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)\mathbb{E}\left(\frac{1}{Y_\infty}\right) - \left(\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)\right)^2 \geqslant 0,$$

where equality would hold if and only if $\mathbb{P}(X_\infty^2 = c) = 1$ with some constant $c \in \mathbb{R}_+$.

Let us use Theorem 2.6 with the following choices

$$d := 2, \qquad M_t := \begin{pmatrix} \int_0^t \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s \\ \int_0^t \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s \end{pmatrix}, \quad t \geqslant 0,$$

$$(\mathcal{F}_t)_{t\geqslant 0} \text{ defined before Proposition 2.1,} \qquad Q(t) := \begin{pmatrix} \frac{1}{\sqrt{t}} & 0 \\ 0 & \frac{1}{\sqrt{t}} \end{pmatrix}, \qquad t > 0.$$

Then

$$\langle M\rangle_t = \begin{pmatrix} \int_0^t \frac{X_s^2}{Y_s}\,\mathrm{d}s & \int_0^t \frac{X_s}{Y_s}\,\mathrm{d}s \\ \int_0^t \frac{X_s}{Y_s}\,\mathrm{d}s & \int_0^t \frac{1}{Y_s}\,\mathrm{d}s \end{pmatrix}, \qquad t \geqslant 0.$$

Recalling that (under the conditions of the theorem) in the proof of Theorem 5.2 it was shown that $\mathbb{E}(1/Y_\infty) < \infty$, $\mathbb{E}(X_\infty/Y_\infty) < \infty$, and $\mathbb{E}(X_\infty^2/Y_\infty) < \infty$, and, hence by Theorem 2.3, we have

$$Q(t)\langle M\rangle_t Q(t)^\top = \begin{pmatrix} \frac{1}{t}\int_0^t \frac{X_s^2}{Y_s}\,\mathrm{d}s & \frac{1}{t}\int_0^t \frac{X_s}{Y_s}\,\mathrm{d}s \\ \frac{1}{t}\int_0^t \frac{X_s}{Y_s}\,\mathrm{d}s & \frac{1}{t}\int_0^t \frac{1}{Y_s}\,\mathrm{d}s \end{pmatrix} \to \begin{pmatrix} \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{1}{Y_\infty}\right) \end{pmatrix} \qquad \text{as} \quad t \to \infty \quad \text{a.s.}$$

Hence, Theorem 2.6 yields (7.2). Then Slutsky's lemma and the continuous mapping theorem yield that

$$\sqrt{T}\begin{pmatrix} \widehat{\theta}_T^{\mathrm{MLE}} - \theta \\ \widehat{m}_T^{\mathrm{MLE}} - m \end{pmatrix} \xrightarrow{\mathcal{L}} \frac{1}{\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)\mathbb{E}\left(\frac{1}{Y_\infty}\right) - \left(\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)\right)^2} \begin{pmatrix} -\mathbb{E}\left(\frac{1}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ -\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) \end{pmatrix} \eta Z \qquad \text{as} \quad T \to \infty.$$

Using that ηZ is a 2-dimensional normally distributed random variable with mean vector zero and with covariance matrix

$$\eta\eta^\top = \begin{pmatrix} \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{1}{Y_\infty}\right) \end{pmatrix},$$

the covariance matrix of

$$\begin{pmatrix} -\mathbb{E}\left(\frac{1}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ -\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) \end{pmatrix} \eta Z$$

takes the form

$$\begin{aligned}
&\begin{pmatrix} -\mathbb{E}\left(\frac{1}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ -\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) \end{pmatrix} \eta\, \mathbb{E}(ZZ^\top)\eta^\top \begin{pmatrix} -\mathbb{E}\left(\frac{1}{Y_\infty}\right) & -\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) \end{pmatrix} \\
&= \begin{pmatrix} -\mathbb{E}\left(\frac{1}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ -\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) \end{pmatrix} \begin{pmatrix} \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{1}{Y_\infty}\right) \end{pmatrix} \begin{pmatrix} -\mathbb{E}\left(\frac{1}{Y_\infty}\right) & -\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) \end{pmatrix} \\
&= \left(\mathbb{E}\left(\frac{1}{Y_\infty}\right)\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) - \left(\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)\right)^2\right) \begin{pmatrix} \mathbb{E}\left(\frac{1}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) \end{pmatrix},
\end{aligned}$$

which yields the assertion. □

7.3 Remark. The asymptotic variance $1/\mathbb{E}\left(X_\infty^2/Y_\infty\right)$ of $\widetilde{\theta}_T^{\mathrm{MLE}}$ in Theorem 7.1 is less than the asymptotic variance

$$\frac{\mathbb{E}\left(\frac{1}{Y_\infty}\right)}{\mathbb{E}\left(\frac{1}{Y_\infty}\right)\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) - \left(\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)\right)^2}$$

of $\widehat{\theta}_T^{\mathrm{MLE}}$ in Theorem 7.2. This is in accordance with the fact that in Theorem 7.1 we supposed that the value of the parameter m is known, which gives extra information, so the MLE estimator of θ becomes better. □

8 Asymptotic behaviour of least squares estimator

8.1 Theorem. *Let us suppose that $m \in \mathbb{R}$ is known. If $a > 0$, $b > 0$, and $\theta > 0$, then the LSE of θ is asymptotically normal, i.e.,*

$$\sqrt{T}(\widetilde{\theta}_T^{\mathrm{LSE}} - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\mathbb{E}(X_\infty^2 Y_\infty)}{(\mathbb{E}(X_\infty^2))^2}\right) \qquad \text{as } T \to \infty,$$

where $\mathbb{E}(X_\infty^2 Y_\infty)$ and $\mathbb{E}(X_\infty^2)$ are given explicitly in Theorem 2.3.

Proof. First note that, by (4.3),

$$\sqrt{T}(\widetilde{\theta}_T^{\mathrm{LSE}} - \theta) = -\frac{\frac{1}{\sqrt{T}}\int_0^T X_s\sqrt{Y_s}\,\mathrm{d}B_s}{\frac{1}{T}\int_0^T X_s^2\,\mathrm{d}s}, \qquad T > 0, \tag{8.1}$$

provided that $\int_0^T X_s^2 \, \mathrm{d}s > 0$. By Theorem 2.3, we have

$$\mathbb{P}\left(\lim_{T\to\infty} \frac{1}{T} \int_0^T X_s^2 \, \mathrm{d}s = \mathbb{E}(X_\infty^2)\right) = 1,$$

$$\mathbb{P}\left(\lim_{T\to\infty} \frac{1}{T} \int_0^T X_s^2 Y_s \, \mathrm{d}s = \mathbb{E}(X_\infty^2 Y_\infty)\right) = 1,$$

where $\mathbb{E}(X_\infty^2)$ and $\mathbb{E}(X_\infty^2 Y_\infty)$ are given explicitly in Theorem 2.3. Note also that $\mathbb{E}(X_\infty^2 Y_\infty)$ is positive (since $X_\infty^2 Y_\infty$ is non-negative and absolutely continuous) and hence

$$\mathbb{P}\left(\lim_{T\to\infty} \int_0^T X_s^2 Y_s \, \mathrm{d}s = +\infty\right) = 1.$$

Further, using Theorem 2.6 with the following choices

$$d := 1, \qquad M_t := \int_0^t X_s \sqrt{Y_s} \, \mathrm{d}B_s, \quad t \geqslant 0, \qquad (\mathcal{F}_t)_{t\geqslant 0} \text{ defined before Proposition 2.1,}$$

$$Q(t) := 1/\sqrt{t}, \quad t > 0, \qquad \eta := \sqrt{\mathbb{E}(X_\infty^2 Y_\infty)},$$

we have

$$\frac{1}{\sqrt{T}} \int_0^T X_s \sqrt{Y_s} \, \mathrm{d}B_s \stackrel{\mathcal{L}}{\longrightarrow} \sqrt{\mathbb{E}(X_\infty^2 Y_\infty)}\,\xi \qquad \text{as } T \to \infty,$$

where ξ is a standard normally distributed random variable. Hence Slutsky's lemma and (8.1) yield the assertion. □

8.2 Remark. The asymptotic variance $\mathbb{E}(X_\infty^2 Y_\infty)/(\mathbb{E}(X_\infty^2))^2$ of the least squares estimator $\widetilde{\theta}_T^{\mathrm{LSE}}$ in Theorem 8.1 is greater than the asymptotic variance $1/\mathbb{E}\left(X_\infty^2/Y_\infty\right)$ of the maximum likelihood estimator $\widetilde{\theta}_T^{\mathrm{MLE}}$ in Theorem 7.1, since, by Cauchy and Schwarz's inequality,

$$(\mathbb{E}(X_\infty^2))^2 = \left(\mathbb{E}\left(\frac{X_\infty}{\sqrt{Y_\infty}} X_\infty \sqrt{Y_\infty}\right)\right)^2 < \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) \mathbb{E}\left(X_\infty^2 Y_\infty\right).$$

□

8.3 Theorem. *If $a > 0$, $b > 0$, $\theta > 0$, and $m \in \mathbb{R}$, then the LSE of (θ, m) is asymptotically normal, i.e.,*

$$\sqrt{T}\begin{pmatrix} \widehat{\theta}_T^{\mathrm{LSE}} - \theta \\ \widehat{m}_T^{\mathrm{LSE}} - m \end{pmatrix} \stackrel{\mathcal{L}}{\longrightarrow} \mathcal{N}_2\left(0, D^{\mathrm{LSE}}\right) \qquad \text{as } T \to \infty,$$

where $\mathcal{N}_2\left(0, D^{\mathrm{LSE}}\right)$ denotes a 2-dimensional normally distribution with mean vector $0 \in \mathbb{R}^2$ and with covariance matrix $D^{\mathrm{LSE}} = (D_{i,j}^{\mathrm{LSE}})_{i,j=1}^2$, where

$$D_{1,1}^{\mathrm{LSE}} := \frac{\mathbb{E}(X_\infty^2 Y_\infty) - 2\,\mathbb{E}(X_\infty)\,\mathbb{E}(X_\infty Y_\infty) + (\mathbb{E}(X_\infty))^2\,\mathbb{E}(Y_\infty)}{(\mathbb{E}(X_\infty^2) - (\mathbb{E}(X_\infty))^2)^2},$$

$$D_{1,2}^{\mathrm{LSE}} = D_{2,1}^{\mathrm{LSE}} := \frac{\mathbb{E}(X_\infty)\big(\mathbb{E}(X_\infty^2 Y_\infty) + \mathbb{E}(X_\infty^2)\,\mathbb{E}(Y_\infty)\big) - \mathbb{E}(X_\infty Y_\infty)\big(\mathbb{E}(X_\infty^2) + (\mathbb{E}(X_\infty))^2\big)}{(\mathbb{E}(X_\infty^2) - (\mathbb{E}(X_\infty))^2)^2},$$

$$D_{2,2}^{\mathrm{LSE}} := \frac{(\mathbb{E}(X_\infty))^2\,\mathbb{E}(X_\infty^2 Y_\infty) - 2\,\mathbb{E}(X_\infty)\,\mathbb{E}(X_\infty^2)\,\mathbb{E}(X_\infty Y_\infty) + (\mathbb{E}(X_\infty^2))^2\,\mathbb{E}(Y_\infty)}{(\mathbb{E}(X_\infty^2) - (\mathbb{E}(X_\infty))^2)^2}.$$

Proof. By (4.8) and (4.9), we have

$$\sqrt{T}(\widehat{\theta}_T^{\mathrm{LSE}} - \theta) = \frac{-\frac{1}{\sqrt{T}}\int_0^T X_s\sqrt{Y_s}\,\mathrm{d}B_s + \frac{1}{T}\int_0^T X_s\,\mathrm{d}s\frac{1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\,\mathrm{d}B_s}{\frac{1}{T}\int_0^T X_s^2\,\mathrm{d}s - \left(\frac{1}{T}\int_0^T X_s\,\mathrm{d}s\right)^2},$$

and

$$\sqrt{T}(\widehat{m}_T^{\mathrm{LSE}} - m) = \frac{-\frac{1}{T}\int_0^T X_s\,\mathrm{d}s\frac{1}{\sqrt{T}}\int_0^T X_s\sqrt{Y_s}\,\mathrm{d}B_s + \frac{1}{T}\int_0^T X_s^2\,\mathrm{d}s\frac{1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\,\mathrm{d}B_s}{\frac{1}{T}\int_0^T X_s^2\,\mathrm{d}s - \left(\frac{1}{T}\int_0^T X_s\,\mathrm{d}s\right)^2},$$

provided that $T\int_0^T X_s^2\,\mathrm{d}s - \left(\int_0^T X_s\,\mathrm{d}s\right)^2 > 0$. Next we show that

$$\left(\frac{1}{\sqrt{T}}\int_0^T X_s\sqrt{Y_s}\,\mathrm{d}B_s, \frac{1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\,\mathrm{d}B_s\right) \xrightarrow{\mathcal{L}} \left((\eta Z)_1, (\eta Z)_2\right) \quad \text{as} \quad T \to \infty, \tag{8.2}$$

where Z is a 2-dimensional standard normally distributed random variable and η is a non-random 2×2 matrix such that

$$\eta\eta^\top = \begin{pmatrix} \mathbb{E}(X_\infty^2 Y_\infty) & \mathbb{E}(X_\infty Y_\infty) \\ \mathbb{E}(X_\infty Y_\infty) & \mathbb{E}(Y_\infty) \end{pmatrix}.$$

Here the matrix

$$\begin{pmatrix} \mathbb{E}(X_\infty^2 Y_\infty) & \mathbb{E}(X_\infty Y_\infty) \\ \mathbb{E}(X_\infty Y_\infty) & \mathbb{E}(Y_\infty) \end{pmatrix}$$

is positive definite, since its principal minors are positive, and η denotes its unique positive definite square root. Indeed, by the absolutely continuity of (Y_∞, X_∞) (see Theorem 2.3), we have $\mathbb{P}(X_\infty^2 Y_\infty = 0) = 0$ and, by Cauchy-Schwarz's inequality,

$$\mathbb{E}(X_\infty^2 Y_\infty)\,\mathbb{E}(Y_\infty) - (\mathbb{E}(X_\infty Y_\infty))^2 \geqslant 0,$$

where equality would hold if and only if $\mathbb{P}(KX_\infty^2 Y_\infty = LY_\infty) = 1$ with some constant $K, L \in \mathbb{R}_+$ such that $K^2 + L^2 > 0$ or equivalently (using that $\mathbb{P}(Y_\infty > 0) = 1$ since Y_∞ has Gamma distribution) if and only if $\mathbb{P}(X_\infty^2 = L/K) = 1$, which leads us to a contradiction refereeing to the absolutely continuity of X_∞.

Let us use Theorem 2.6 with the following choices

$$d := 2, \qquad M_t := \begin{pmatrix} \int_0^t X_s\sqrt{Y_s}\,\mathrm{d}B_s \\ \int_0^t \sqrt{Y_s}\,\mathrm{d}B_s \end{pmatrix}, \quad t \geqslant 0,$$

$$(\mathcal{F}_t)_{t\geqslant 0} \text{ defined before Proposition 2.1,} \qquad Q(t) := \begin{pmatrix} \frac{1}{\sqrt{t}} & 0 \\ 0 & \frac{1}{\sqrt{t}} \end{pmatrix}, \quad t > 0.$$

Then

$$\langle M\rangle_t = \begin{pmatrix} \int_0^t X_s^2 Y_s\,\mathrm{d}s & \int_0^t X_s Y_s\,\mathrm{d}s \\ \int_0^t X_s Y_s\,\mathrm{d}s & \int_0^t Y_s\,\mathrm{d}s \end{pmatrix}, \qquad t \geqslant 0,$$

and hence, by Theorem 2.3 (similarly as detailed in the proof of Theorem 7.2),

$$Q(t)\langle M\rangle_t Q(t)^\top = \begin{pmatrix} \frac{1}{t}\int_0^t X_s^2 Y_s \,\mathrm{d}s & \frac{1}{t}\int_0^t X_s Y_s \,\mathrm{d}s \\ \frac{1}{t}\int_0^t X_s Y_s \,\mathrm{d}s & \frac{1}{t}\int_0^t Y_s \,\mathrm{d}s \end{pmatrix} \to \begin{pmatrix} \mathbb{E}(X_\infty^2 Y_\infty) & \mathbb{E}(X_\infty Y_\infty) \\ \mathbb{E}(X_\infty Y_\infty) & \mathbb{E}(Y_\infty) \end{pmatrix} \qquad \text{as} \quad t \to \infty \quad \text{a.s.}$$

Note also that, by Theorem 2.3, we have

$$\mathbb{P}\left(\lim_{T\to\infty} \frac{1}{T} \int_0^T X_s \,\mathrm{d}s = \mathbb{E}(X_\infty) \right) = 1 \qquad \text{and} \qquad \mathbb{P}\left(\lim_{T\to\infty} \frac{1}{T} \int_0^T X_s^2 \,\mathrm{d}s = \mathbb{E}(X_\infty^2) \right) = 1.$$

Hence Slutsky's lemma and the continuous mapping theorem yield that

$$\sqrt{T} \begin{pmatrix} \widehat{\theta}_T^{\mathrm{LSE}} - \theta \\ \widehat{m}_T^{\mathrm{LSE}} - \theta \end{pmatrix} \xrightarrow{\mathcal{L}} \frac{1}{\mathbb{E}(X_\infty^2) - (\mathbb{E}(X_\infty))^2} \begin{pmatrix} -1 & \mathbb{E}(X_\infty) \\ -\mathbb{E}(X_\infty) & \mathbb{E}(X_\infty^2) \end{pmatrix} \eta Z \qquad \text{as} \quad T \to \infty.$$

Using that ηZ is a 2-dimensional normally distributed random variable with mean vector zero and with covariance matrix

$$\eta\eta^\top = \begin{pmatrix} \mathbb{E}(X_\infty^2 Y_\infty) & \mathbb{E}(X_\infty Y_\infty) \\ \mathbb{E}(X_\infty Y_\infty) & \mathbb{E}(Y_\infty) \end{pmatrix},$$

the covariance matrix of

$$\begin{pmatrix} -1 & \mathbb{E}(X_\infty) \\ -\mathbb{E}(X_\infty) & \mathbb{E}(X_\infty^2) \end{pmatrix} \eta Z$$

takes the form

$$\begin{aligned}
&\begin{pmatrix} -1 & \mathbb{E}(X_\infty) \\ -\mathbb{E}(X_\infty) & \mathbb{E}(X_\infty^2) \end{pmatrix} \eta\,\mathbb{E}(ZZ^\top)\eta^\top \begin{pmatrix} -1 & -\mathbb{E}(X_\infty) \\ \mathbb{E}(X_\infty) & \mathbb{E}(X_\infty^2) \end{pmatrix} \\
&= \begin{pmatrix} -1 & \mathbb{E}(X_\infty) \\ -\mathbb{E}(X_\infty) & \mathbb{E}(X_\infty^2) \end{pmatrix} \begin{pmatrix} \mathbb{E}(X_\infty^2 Y_\infty) & \mathbb{E}(X_\infty Y_\infty) \\ \mathbb{E}(X_\infty Y_\infty) & \mathbb{E}(Y_\infty) \end{pmatrix} \begin{pmatrix} -1 & -\mathbb{E}(X_\infty) \\ \mathbb{E}(X_\infty) & \mathbb{E}(X_\infty^2) \end{pmatrix} \\
&= \begin{pmatrix} -\mathbb{E}(X_\infty^2 Y_\infty) + \mathbb{E}(X_\infty)\,\mathbb{E}(X_\infty Y_\infty) & -\mathbb{E}(X_\infty Y_\infty) + \mathbb{E}(X_\infty)\,\mathbb{E}(Y_\infty) \\ -\mathbb{E}(X_\infty)\,\mathbb{E}(X_\infty^2 Y_\infty) + \mathbb{E}(X_\infty^2)\,\mathbb{E}(X_\infty Y_\infty) & -\mathbb{E}(X_\infty)\,\mathbb{E}(X_\infty Y_\infty) + \mathbb{E}(X_\infty^2)\,\mathbb{E}(Y_\infty) \end{pmatrix} \\
&\quad \times \begin{pmatrix} -1 & -\mathbb{E}(X_\infty) \\ \mathbb{E}(X_\infty) & \mathbb{E}(X_\infty^2) \end{pmatrix},
\end{aligned}$$

which yields the assertion. □

8.4 Remark. Using the explicit forms of the mixed moments given in (iii) of Theorem 2.3, one can check that the asymptotic variance $\mathbb{E}(X_\infty^2 Y_\infty)/(\mathbb{E}(X_\infty^2))^2$ of $\widetilde{\theta}_T^{\mathrm{LSE}}$ in Theorem 8.1 is less than the asymptotic variance $D_{1,1}^{\mathrm{LSE}}$ of $\widehat{\theta}_T^{\mathrm{LSE}}$ in Theorem 8.3. This can be interpreted similarly as in Remark 7.3. □

References

[1] Barczy, M., Döring, L., Li, Z. and Pap, G. (2012). On parameter estimation for critical affine processes. Available on ArXiv: `http://www.arxiv.org/abs/1210.1866`

[2] Barczy, M., Döring, L., Li, Z. and Pap, G. (2013). Ergodicity for an affine two factor model. Available on ArXiv: `http://arxiv.org/abs/1302.2534`

[3] Barndorff-Nielsen, O. and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B* **63** 167–241.

[4] Ben Alaya, M. and Kebaier, M. (2012). Parameter estimation for the square root diffusions: ergodic and nonergodic cases. *Stochastic Models* **28**(4) 609–634.

[5] Ben Alaya, M. and Kebaier, M. (2012). Asymptotic behavior of the maximum likelihood estimator for ergodic and nonergodic square-root diffusions. Available on the HAL: `http://hal.archives-ouvertes.fr/hal-00640053`

[6] Bishwal, J. P. N. (2008). *Parameter Estimation in Stochastic Differential Equations*. Springer-Verlag Berlin Heidelberg.

[7] Carr, P. and Wu, L. (2003). The finite moment log stable process and option pricing. *The Journal of Finance* **58** 753–777.

[8] Chen, H. and Joslin, S. (2012). Generalized transform analysis of affine processes and applications in finance. *Review of Financial Studies* **25**(7) 2225–2256.

[9] Dawson, D. A. and Li, Z. (2006). Skew convolution semigroups and affine Markov processes. *The Annals of Probability* **34(3)** 1103–1142.

[10] Duffie, D., Filipović, D. and Schachermayer, W. (2003). Affine processes and applications in finance. *Annals of Applied Probability* **13** 984–1053.

[11] Heston, S. (1993). A closed-form solution for options with stochastic volatilities with applications to bond and currency options. *The Review of Financial Studies* **6** 327–343.

[12] Hu, Y. and Long, H. (2007). Parameter estimation for Ornstein–Uhlenbeck processes driven by α-stable Lévy motions. *Communications on Stochastic Analysis* **1(2)** 175–192.

[13] Hu, Y. and Long, H. (2009). Least squares estimator for Ornstein–Uhlenbeck processes driven by α-stable motions. *Stochastic Processes and their Applications* **119(8)** 2465–2480.

[14] Hu, Y. and Long, H. (2009). On the singularity of least squares estimator for mean-reverting α-stable motions. *Acta Mathematica Scientia* **29B(3)** 599–608.

[15] Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. Springer-Verlag, Berlin.

[16] Kutoyants, Yu A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. Springer-Verlag London Limited.

[17] Li, Z. (2011). *Measure-Valued Branching Markov Processes*. Springer-Verlag, Heidelberg.

[18] Li, Z. and Ma, C. (2013). Asymptotic properties of estimators in a stable Cox-Ingersoll-Ross model. Available on the ArXiv: `http://arxiv.org/abs/1301.3243`

[19] Liptser, R. S. and Shiryaev, A. N. (2001). *Statistics of Random Processes I. Applications*, 2nd edition. Springer-Verlag, Berlin, Heidelberg.

[20] Liptser, R. S. and Shiryaev, A. N. (2001). *Statistics of Random Processes II. Applications*, 2nd edition. Springer-Verlag, Berlin, Heidelberg.

[21] Overbeck, L. and Rydén, T. (1997). Estimation in the Cox-Ingersoll-Ross model. *Econometric Theory* **13**(3) 430–461.

[22] Overbeck, L. (1998). Estimation for continuous branching processes. *Scandinavian Journal of Statistics* **25**(1) 111–126.

[23] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. Springer-Verlag Berlin Heidelberg.

[24] Sandrić, N. (2013). Long-time behaviour of stable-like processes. *Stochastic Processes and their Applications* **123**(4) 1276–1300.

[25] van Zanten, H. (2000). A multivariate central limit theorem for continuous local martingales. *Statistics & Probability Letters* **50**(3) 229–235.

Mátyás Barczy, Faculty of Informatics, University of Debrecen, Pf. 12, H–4010 Debrecen, Hungary. Tel.: +36-52-512900, Fax: +36-52-512996, e–mail: barczy.matyas@inf.unideb.hu

Leif Döring, Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland. Tel.: +41-(0)44-63 55892, Fax: +41-(0)44-63 55705,
e–mail: leif.doering@googlemail.com

Zenghu Li, School of Mathematical Sciences, Beijing Normal University, Beijing 100875, People's Republic of China. Tel.: +86-10-58802900, Fax: +86-10-58808202, e–mail: lizh@bnu.edu.cn

Gyula Pap, Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary. Tel.: +36-62-544033, Fax: +36-62-544548, e–mail: papgy@math.u-szeged.hu